

NONUNIFORM SAMPLING, REPRODUCING KERNELS, AND THE ASSOCIATED HILBERT SPACES

PALLE JORGENSEN AND FENG TIAN

ABSTRACT. In a general context of positive definite kernels k , we develop tools and algorithms for sampling in reproducing kernel Hilbert space \mathcal{H} (RKHS). With reference to these RKHSs, our results allow inference from samples; more precisely, reconstruction of an “entire” (or global) signal, a function f from \mathcal{H} , via generalized interpolation of f from partial information obtained from carefully chosen distributions of sample points. We give necessary and sufficient conditions for configurations of point-masses δ_x of sample-points x to have finite norm relative to the particular RKHS \mathcal{H} considered. When this is the case, and the kernel k is given, we obtain an induced positive definite kernel $\langle \delta_x, \delta_y \rangle_{\mathcal{H}}$. We perform a comparison, and we study when this induced positive definite kernel has l^2 rows and columns. The latter task is accomplished with the use of certain symmetric pairs of operators in the two Hilbert spaces, l^2 on one side, and the RKHS \mathcal{H} on the other. A number of applications are given, including to infinite network systems, to graph Laplacians, to resistance metrics, and to sampling of Gaussian fields.

CONTENTS

1. Introduction	2
2. Reproducing kernel Hilbert spaces (RKHSs), and relative RKHSs	3
3. Sampling and point-masses of finite norm	6
4. A symmetric pair of operators associated with a RKHS having its point-masses of finite norm	8
4.1. Application: Moment analysis of networks with given conductance function	17
4.2. Discrete sample points for Brownian motion	18
5. Spectral theory: A necessary and sufficient condition for when the symmetric pair is maximal	18
6. Sample point-masses in concrete models	20
6.1. Sample points in Brownian motion	21
6.2. Discrete RKHSs from restrictions	24
6.3. Brownian bridge	26
6.4. Binomial RKHS	27
6.5. Classical RKHSs with point-mass samples and interpolation	29

2000 *Mathematics Subject Classification.* Primary 47L60, 46N30, 46N50, 42C15, 65R10, 31C20, 62D05, 94A20, 39A12; Secondary 46N20, 22E70, 31A15, 58J65.

Key words and phrases. Computational harmonic analysis, Hilbert space, reproducing kernel Hilbert space, unbounded operators, discrete analysis, interpolation, reconstruction, infinite matrices, binomial coefficients, Gaussian free fields, graph Laplacians, distribution of point-masses, Green’s function (graph Laplacians), non-uniform sampling, transforms, (discrete) Ito-isometries, optimization, infinite determinants, moments, covariance, interpolation.

1. INTRODUCTION

In the theory of non-uniform sampling, one studies Hilbert spaces consisting of signals, understood in a very general sense. One then develops analytic tools and algorithms, allowing one to draw inference for an “entire” (or global) signal from partial information obtained from carefully chosen distributions of sample points. While the better known and classical sampling algorithms (Shannon and others) are based on interpolation, modern theories go beyond this. An early motivation is the work of Henry Landau. In this setting, it is possible to make precise the notion of “average sampling rates” in general configurations of sample points. Our present study, turns the tables. We start with the general axiom system of positive definite kernels and their associated reproducing kernel Hilbert spaces (RKHSs), or *relative* RKHSs. With some use of metric geometry and of spectral theory for operators in Hilbert space, we are then able to obtain sampling theorems for a host of non-uniform point configurations. The modern theory of non-uniform sampling is vast, and it specializes into a variety of sub-areas. The following papers (and the literature cited there) will give an idea of the diversity of points of view: [KGD13, ZLP14, MGG14, AL08, Lan67].

A reproducing kernel Hilbert space (RKHS) is a Hilbert space \mathcal{H} of functions on a prescribed set, say V , with the property that point-evaluation for functions $f \in \mathcal{H}$ is continuous with respect to the \mathcal{H} -norm. They are called kernel spaces, because, for every $x \in V$, the point-evaluation for functions $f \in \mathcal{H}$, $f(x)$ must then be given as a \mathcal{H} -inner product of f and a vector k_x in \mathcal{H} ; called the kernel, i.e., $f(x) = \langle k_x, f \rangle_{\mathcal{H}}$, $\forall f \in \mathcal{H}$, $x \in V$.

There is a related reproducing kernel notion called “relative.” This means that increments have kernel representations. In detail: Consider functions f in \mathcal{H} , but suppose instead that, for every pair of points x, y in V , each of the differences $f(x) - f(y)$ can be represented by a kernel from \mathcal{H} . We then say that \mathcal{H} is a relative RKHS. We shall study both in our paper. The “relative” variant is of more recent vintage, and it is used in the study of electrical networks (voltage differences, see Lemma 3.3); and in analysis of Gaussian processes such as Gaussian fields, sect 6.1 and 6.3.

The RKHSs have been studied extensively since the pioneering papers by Aronszajn [Aro43, Aro48]. They further play an important role in the theory of partial differential operators (PDO); for example as Green’s functions of second order elliptic PDOs [Nel57, HKL⁺14]. Other applications include engineering, physics, machine-learning theory [KH11, SZ09, CS02], stochastic processes [AD93, ABDdS93, AD92, AJSV13, AJV14], numerical analysis, and more [LB04, HQKL10, ZXZ12, LP11, CFM⁺13, Vul13, SS13, HN14, STC04, SS01]. But the literature so far has focused on the theory of kernel functions defined on continuous domains, either domains in Euclidean space, or complex domains in one or more variables. For these cases, the Dirac δ_x distributions do not have finite \mathcal{H} -norm. But for RKHSs over discrete point distributions, it is reasonable to expect that the Dirac δ_x functions will in fact have finite \mathcal{H} -norm.

An illustration from *neural networks*: An Extreme Learning Machine (ELM) is a neural network configuration in which a hidden layer of weights are randomly sampled [RW06], and the object is then to determine analytically resulting output layer weights. Hence ELM may be thought of as an approximation to a network with infinite number of hidden units.

The main results in our paper include Theorem 2.10, Corollary 2.11, Theorem 4.8, Theorem 4.10, Corollary 4.15, and Theorem 4.20 where we give necessary and sufficient conditions for the point-masses to have finite norm relative to the particular RKHS \mathcal{H} considered. When this is the case, we obtain an induced positive definite kernel $\langle \delta_x, \delta_y \rangle_{\mathcal{H}}$. In Section 4, we study when this induced positive definite kernel has l^2 rows and columns. The latter task is accomplished with the use of certain symmetric pairs of operators in the two Hilbert spaces, l^2 on one side, and the RKHS \mathcal{H} on the other. In Section 5, we study the cases when the associated symmetric pair is maximal. The results from Sections 4-5 are then applied in Section 6 to the study of admissible distributions of discrete sample points for Brownian motion, and for related Gaussian fields. We have a separate subsection 6.4 discussing the RKHS constructed canonically from the binomial coefficients.

2. REPRODUCING KERNEL HILBERT SPACES (RKHSs), AND RELATIVE RKHSs

Here we consider *the discrete case*, i.e., RKHSs of functions defined on a prescribed countable infinite discrete set V . We are concerned with a characterization of those RKHSs \mathcal{H} which contain the Dirac masses δ_x for all points $x \in V$. Of the examples and applications where this question plays an important role, we emphasize three: (i) *discrete Brownian motion-Hilbert spaces*, i.e., discrete versions of the Cameron-Martin Hilbert space; (ii) *energy-Hilbert spaces* corresponding to graph-Laplacians; and finally (iii) RKHSs generated by *binomial coefficients*.

In general when reproducing kernels and their Hilbert spaces are used, one ends up with functions on a suitable set, and so far we feel that the dichotomy *discrete* vs *continuous* has not yet received sufficient attention. After all, a choice of sampling points in relevant optimization models based on kernel theory suggests the need for a better understanding of point masses as they are accounted for in the RKHS at hand. In broad outline, this is a leading theme in our paper.

The two definitions below, and Lemma 2.4 are valid more generally for the setting when V is an arbitrary set. But we have nonetheless restricted our focus to the case when V is assumed countably infinite. The reason for this will become evident in Definition 2.5, and in Lemma 2.8, Corollary 2.9, and Theorem 2.10, to follow.

Definition 2.1. Let V be a countable and infinite set, and $\mathcal{F}(V)$ the set of all *finite* subsets of V . A function $k : V \times V \rightarrow \mathbb{C}$ is said to be *positive definite*, if

$$\sum_{(x,y) \in F \times F} k(x,y) \overline{c_x} c_y \geq 0 \quad (2.1)$$

holds for all coefficients $\{c_x\}_{x \in F} \subset \mathbb{C}$, and all $F \in \mathcal{F}(V)$.

Definition 2.2. Fix a set V , countable infinite.

- (1) For all $x \in V$, set

$$k_x := k(\cdot, x) : V \rightarrow \mathbb{C} \quad (2.2)$$

as a function on V .

- (2) Let $\mathcal{H} := \mathcal{H}(k)$ be the Hilbert-completion of the $\text{span}\{k_x \mid x \in V\}$, with respect to the inner product

$$\left\langle \sum c_x k_x, \sum d_y k_y \right\rangle_{\mathcal{H}} := \sum \sum \overline{c_x} d_y k(x, y) \quad (2.3)$$

modulo the subspace of functions of zero \mathcal{H} -norm. \mathcal{H} is then a reproducing kernel Hilbert space (RKHS), with the reproducing property:

$$\langle k_x, \varphi \rangle_{\mathcal{H}} = \varphi(x), \quad \forall x \in V, \quad \forall \varphi \in \mathcal{H}. \quad (2.4)$$

- (3) If $F \in \mathcal{F}(V)$, set $\mathcal{H}_F = \text{closed span}\{k_x\}_{x \in F} \subset \mathcal{H}$, (closed is automatic if F is finite.) And set

$$P_F := \text{the orthogonal projection onto } \mathcal{H}_F. \quad (2.5)$$

- (4) For $F \in \mathcal{F}(V)$, set

$$K_F := (k(x, y))_{(x, y) \in F \times F} \quad (2.6)$$

as a $\#F \times \#F$ matrix.

Remark 2.3. The summations in (2.3) are all finite. Starting with finitely supported summations in (2.3), the RKHS $\mathcal{H} (= \mathcal{H}(k))$ is then obtained by Hilbert space completion. We use physicists' convention, so that the inner product is conjugate linear in the first variable, and linear in the second variable.

The following result is known; and it follows from the definitions above.

Lemma 2.4. *Let $k : V \times V \rightarrow \mathbb{C}$ be positive definite, and let \mathcal{H} be the corresponding RKHS. Let f be an arbitrary function on V . Then f is in \mathcal{H} if and only if there is a constant $C = C_f < \infty$ such that, for every finitely supported function $\xi : V \rightarrow \mathbb{C}$, we have the estimate*

$$\left| \sum_{x \in V} \overline{\xi(x)} f(x) \right|^2 \leq C \sum_{x, y \in V} \overline{\xi(x)} \xi(y) k(x, y)$$

with the constant $C = C_f$ independent of ξ .

It follows from the above that reproducing kernel Hilbert spaces (RKHSs) arise from a given positive definite kernel k , a corresponding pre-Hilbert form; and then a Hilbert-completion. The question arises: "What are the functions in the completion?" Now, before completion, the functions are as specified in Definition 2.2, but the Hilbert space completions are subtle; they are classical Hilbert spaces of functions, not always transparent from the naked kernel k itself. Examples of classical RKHSs: Hardy spaces or Bergman spaces (for complex domains), Sobolev spaces and Dirichlet spaces [OST13, ST12, Str10] (for real domains, or for fractals), band-limited L^2 functions (from signal analysis), and Cameron-Martin Hilbert spaces from Gaussian processes (in continuous time domain).

Our focus here is on discrete analogues of the classical RKHSs from real or complex analysis. These discrete RKHSs in turn are dictated by applications, and their features are quite different from those of their continuous counterparts.

Definition 2.5. The RKHS $\mathcal{H} = \mathcal{H}(k)$ is said to have the *discrete mass* property (\mathcal{H} is called a *discrete RKHS*), if $\delta_x \in \mathcal{H}$, for all $x \in V$. Here, $\delta_x(y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$, i.e., the Dirac mass at $x \in V$.

Lemma 2.6. *Let $F \in \mathcal{F}(V)$, $x_1 \in F$. Assume $\delta_{x_1} \in \mathcal{H}$. Then*

$$P_F(\delta_{x_1})(\cdot) = \sum_{y \in F} (K_F^{-1} \delta_{x_1})(y) k_y(\cdot). \quad (2.7)$$

Proof. We check that

$$\delta_{x_1} - \sum_{y \in F} (K_F^{-1} \delta_{x_1})(y) k_y(\cdot) \in \mathcal{H}_F^\perp. \quad (2.8)$$

The remaining part follows easily from this.

(The notation $(\mathcal{H}_F)^\perp$ stands for orthogonal complement, also denoted $\mathcal{H} \ominus \mathcal{H}_F = \{\varphi \in \mathcal{H} \mid \langle f, \varphi \rangle_{\mathcal{H}} = 0, \forall f \in \mathcal{H}_F\}$.) \square

Remark 2.7. A slight abuse of notations: We make formally sense of the expressions for $P_F(\delta_x)$ in (2.7) even in the case when δ_x might not be in \mathcal{H} . For all finite F , we showed that $P_F(\delta_x) \in \mathcal{H}$. But for δ_x be in \mathcal{H} , we must have the additional boundedness assumption (2.14) satisfied; see Theorem 2.10.

Lemma 2.8. *Let $F \in \mathcal{F}(V)$, $x_1 \in F$, then*

$$(K_F^{-1} \delta_{x_1})(x_1) = \|P_F \delta_{x_1}\|_{\mathcal{H}}^2. \quad (2.9)$$

Proof. Setting $\zeta^{(F)} := K_F^{-1}(\delta_{x_1})$, we have

$$P_F(\delta_{x_1}) = \sum_{y \in F} \zeta^{(F)}(y) k_F(\cdot, y)$$

and for all $z \in F$,

$$\underbrace{\sum_{z \in F} \zeta^{(F)}(z) P_F(\delta_{x_1})(z)}_{\zeta^{(F)}(x_1)} = \sum_F \sum_F \zeta^{(F)}(z) \zeta^{(F)}(y) k_F(z, y) \quad (2.10)$$

$$= \|P_F \delta_{x_1}\|_{\mathcal{H}}^2.$$

By Lemma 2.6, the LHS of (2.10) is given by

$$\begin{aligned} \|P_F \delta_{x_1}\|_{\mathcal{H}}^2 &= \langle P_F \delta_{x_1}, \delta_{x_1} \rangle_{\mathcal{H}} \\ &= \sum_{y \in F} (K_F^{-1} \delta_{x_1})(y) \langle k_y, \delta_{x_1} \rangle_{\mathcal{H}} \\ &= (K_F^{-1} \delta_{x_1})(x_1) = K_F^{-1}(x_1, x_1). \end{aligned}$$

\square

Corollary 2.9. *If $\delta_{x_1} \in \mathcal{H}$ (see Theorem 2.10), then*

$$\sup_{F \in \mathcal{F}(V)} (K_F^{-1} \delta_{x_1})(x_1) = \|\delta_{x_1}\|_{\mathcal{H}}^2. \quad (2.11)$$

Theorem 2.10. *Given V , $k : V \times V \rightarrow \mathbb{R}$ positive definite (p.d.). Let $\mathcal{H} = \mathcal{H}(k)$ be the corresponding RKHS. Assume V is countably infinite. Then the following three conditions (1)-(3) are equivalent; $x_1 \in V$ is fixed:*

- (1) $\delta_{x_1} \in \mathcal{H}$;
- (2) $\exists C_{x_1} < \infty$ such that for all $F \in \mathcal{F}(V)$, the following estimate holds:

$$|\xi(x_1)|^2 \leq C_{x_1} \sum_{(x,y) \in F \times F} \overline{\xi(x)} \xi(y) k(x, y) \quad (2.12)$$

(3) For $F \in \mathcal{F}(V)$, set

$$K_F = (k(x, y))_{(x, y) \in F \times F} \quad (2.13)$$

as a $\#F \times \#F$ matrix. Then

$$\sup_{F \in \mathcal{F}(V)} (K_F^{-1} \delta_{x_1})(x_1) < \infty. \quad (2.14)$$

Proof. For details, see [JT15b, JT15a]. See also Lemma 2.4. \square

Following [KZ96], we say that k is *strictly positive* iff $\det K_F > 0$ for all $F \in \mathcal{F}(V)$.

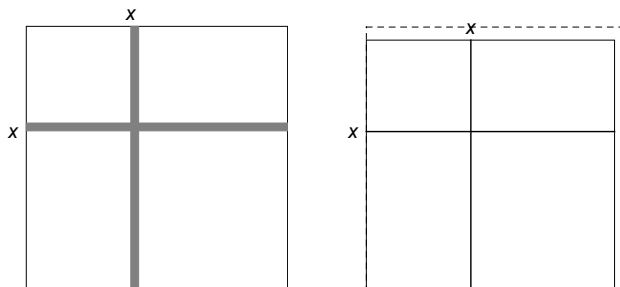


FIGURE 2.1. The (x, x) minors, $K_F \rightarrow K'_F$.

Corollary 2.11. Suppose $k : V \times V \rightarrow \mathbb{R}$ is strictly positive. Set $D_F := \det K_F$. If $x \in V$, and $F \in \mathcal{F}_x(V)$, set $K'_F :=$ the minor in K_F obtained by omitting row x and column x , see Figure 2.1. Let $x \in V$. Then

$$\delta_x \in \mathcal{H} \iff \sup_{F \in \mathcal{F}_x(V)} \frac{D'_F}{D_F} < \infty. \quad (2.15)$$

3. SAMPLING AND POINT-MASSSES OF FINITE NORM

The results presented below hold both for the case of reproducing kernel Hilbert spaces (RKHSs), and the parallel case of *relative* RKHSs. However, we shall state theorems only in the first case. The reader will be able to formulate the results in the case of relative RKHSs. The proofs in the relative case are the same but with slight modifications *mutatis mutandis*.

An important special case of relative RKHSs is that of infinite networks (or graphs) treated in the earlier literature.

Infinite vs finite graphs. We study “large weighted graphs” (vertices V , edges E , and weights as functions assigned on the edges E), and our motivation derives from learning where “learning” is understood broadly to include (machine) learning of suitable probability distribution, i.e., meaning learning from samples of training data. Other applications of an analysis of weighted graphs include statistical mechanics, such as infinite spin models, and large digital networks. It is natural to ask then how one best approaches analysis on “large” systems. We propose an analysis via infinite weighted graphs. This is so even if some of the questions in learning theory may in fact refer to only “large” finite graphs.

One reason for this (among others) is that *statistical features* in such an analysis are best predicted by consideration of probability spaces corresponding to measures on infinite sample spaces. Moreover the latter are best designed from consideration of infinite weighted graphs, as opposed to their finite counterparts. Examples of statistical features which are relevant even for finite samples is long-range order; i.e., the study of correlations between distant sites (vertices), and related phase-transitions, e.g., sign-flips at distant sites. In designing efficient learning models, it is important to understand the possible occurrence of unexpected long-range correlations; e.g., correlations between distant sites in a finite sample.

A second reason for the use of infinite sample-spaces is their use in designing efficient sampling procedures. The interesting solutions will often occur first as vectors in an infinite-dimensional reproducing-kernel Hilbert space RKHS. Indeed, such RKHSs serve as powerful tools in the solution of a kernel-optimization problems with penalty terms. Once an optimal solution is obtained in infinite dimensions, one may then proceed to study its restrictions to suitably chosen finite subgraphs. See [JS09, JS12, JS13, JPT14, JT15c].

Definition 3.1. An infinite network consists of the following:

- V a set of vertices, $\#V = \aleph_0$;
- $E \subset V \times V \setminus \{\text{diagonal}\}$, edges;
- $c : E \rightarrow \mathbb{R}_+$ a fixed symmetric function representing conductance.

We assume (V, E, c) is connected, i.e., for $\forall x, y \in V$, $\exists (x_i x_{i+1})_{i=0}^{n-1} \in E$ s.t. $x_0 = x$, $x_n = y$.

Definition 3.2. Let (V, E, c) be an infinite network. We denote by \mathcal{H}_E the *energy Hilbert space*, where

$$\mathcal{H}_E = \left\{ f : V \rightarrow \mathbb{C} \mid \|f\|_{\mathcal{H}_E}^2 := \frac{1}{2} \sum_{(xy) \in E} \sum_{(xy) \in E} c_{xy} |f(x) - f(y)|^2 < \infty \right\}.$$

Given $f, g \in \mathcal{H}_E$, the inner product is

$$\langle f, g \rangle_{\mathcal{H}_E} = \frac{1}{2} \sum_{(xy) \in E} \sum_{(xy) \in E} c_{xy} (\overline{f(x)} - \overline{f(y)}) (f(x) - f(y)).$$

The following two facts are well-known:

Lemma 3.3. Let V, E, \mathcal{H}_E be as in Definitions 3.1 and 3.2. Then

- (1) \mathcal{H}_E is a Hilbert space; and
- (2) For $\forall x, y \in V$, $\exists! v_{xy} \in \mathcal{H}_E$ (called *dipole*) s.t.

$$f(x) - f(y) = \langle v_{xy}, f \rangle_{\mathcal{H}_E}, \quad \forall f \in \mathcal{H}_E. \quad (3.1)$$

Proof. While this is in the literature, we will include a brief sketch. Part (1) is clear. To prove (2), recall that it is assumed that (V, E) is connected; so given any pair $x, y \in V$, $\exists n \in \mathbb{N}$, and $(x_i x_{i+1})_{i=0}^{n-1} \in E$ s.t. $x_0 = x$ and $x_n = y$. Then, for $\forall f \in \mathcal{H}_E$, we have

$$\begin{aligned} |f(y) - f(x)| &= \left| \sum_{i=0}^{n-1} f(x_{i+1}) - f(x_i) \right| \\ &= \left| \sum_{i=0}^{n-1} \frac{1}{\sqrt{c_{x_i x_{i+1}}}} \sqrt{c_{x_i x_{i+1}}} (f(x_{i+1}) - f(x_i)) \right| \end{aligned}$$

$$\begin{aligned}
(\text{Schwarz}) &\leq \left(\sum_{i=0}^{n-1} \frac{1}{c_{x_i x_{i+1}}} \right)^{\frac{1}{2}} \left(\sum_{i=0}^{n-1} c_{x_i x_{i+1}} |f(x_{i+1}) - f(x_i)|^2 \right)^{\frac{1}{2}} \\
&\leq \text{Const} \cdot \|f\|_{\mathcal{H}_E}.
\end{aligned}$$

The existence of $v_{xy} \in \mathcal{H}_E$ as asserted in (3.1) now follows from an application of Riesz' theorem to the Hilbert space \mathcal{H}_E . Also see [JT15c, JP10, JT15b]. \square

Definition 3.4. It will be convenient to choose a fixed base-point, say $o \in V$, and set $v_x := v_{xo}$. In this case, (3.1) takes the form

$$f(x) - f(o) = \langle v_{xo}, f \rangle_{\mathcal{H}_E}, \quad \forall f \in \mathcal{H}_E. \quad (3.2)$$

We say that \mathcal{H}_E is a relative RKHS. The corresponding positive definite kernel is as follows:

$$k(x, y) = \langle v_x, v_y \rangle_{\mathcal{H}_E}, \quad (x, y) \in V \times V.$$

We say that a given infinite network (V, E, c, \mathcal{H}_E) as above has finite point-masses iff

$$\delta_x \in \mathcal{H}_E, \quad \forall x \in V. \quad (3.3)$$

Remark 3.5. The condition (3.3) will be automatic if for all $x \in V$,

$$\#\{y \in V \mid (xy) \in E\} < \infty, \quad (3.4)$$

but the finite-point mass case holds in many examples where (3.4) is not assumed, see Section 4 below.

Proposition 3.6. Let (V, E, c) and \mathcal{H}_E be as in Definition 3.2. Assume condition (3.4) is satisfied. For functions f on V , set

$$(\Delta_c f)(x) = \sum_{y \sim x} c_{xy} (f(x) - f(y)). \quad (3.5)$$

Finally, let $\{v_x\}_{x \in V \setminus \{o\}}$ be a system of dipoles; see Definition 3.4. Set

$$k_x(y) := \langle v_y, v_x \rangle_{\mathcal{H}_E}, \quad y \in V; \quad (3.6)$$

then

$$(\Delta_c k_x)(y) = \delta_{x,y}, \quad \forall x, y \in V \setminus \{o\}. \quad (3.7)$$

Proof. The verification of (3.7) is a direct computation which we leave to the reader.

Because of (3.7), one often says that k (in (3.6)) is a Green's function for the graph Laplacian Δ_c in (3.5).

For other applications of related semibounded selfadjoint operators, see e.g., [JPT14]. \square

4. A SYMMETRIC PAIR OF OPERATORS ASSOCIATED WITH A RKHS HAVING ITS POINT-MASSSES OF FINITE NORM

We now turn to the general case of positive definite kernels and the case of RKHSs, and *relative* RKHSs, such that the point-mass condition (3.3) is satisfied. We show that there is then an associated and canonical symmetric pair of operators (A, B) :

Definition 4.1. Let $k : V \times V \rightarrow \mathbb{C}$ (or \mathbb{R}) be a positive definite kernel, and \mathcal{H} be the corresponding RKHS as above. If (2.12) holds, i.e., \mathcal{H} has the finite-mass property (Def. 2.5), then we get a dual pair of operators as follows (see Fig 4.1):

$A : l^2(V) \rightarrow \mathcal{H} (= \mathcal{H}(k))$, $\mathcal{D}(A) = \text{span}\{\delta_x\}$ dense in $l^2(V)$, with

$$A\delta_x = \delta_x \in \mathcal{H};$$

$B : \mathcal{H} \rightarrow l^2(V)$, $\mathcal{D}(B) = \text{span}\{k_x\}$ dense in \mathcal{H} , where

Case 1. RKHS:

$$Bk_x = \delta_x \quad (4.1)$$

Case 2. Relative RKHS (Definition 3.4):

$$Bk_x = \delta_x - \delta_o. \quad (4.2)$$

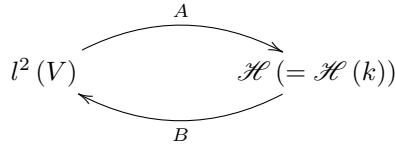


FIGURE 4.1. The pair of operators (A, B) .

Proposition 4.2. The system (A, B) from Definition 4.1 is a symmetric pair, i.e.,

$$\langle Au, v \rangle_{\mathcal{H}} = \langle u, Bv \rangle_{l^2}, \quad \forall u \in \mathcal{D}(A), \forall v \in \mathcal{D}(B). \quad (4.3)$$

Proof. It suffices to consider real Hilbert spaces (the modifications needed for the complex case are straightforward), in which case we have:

$$\begin{aligned} &\langle u, v \rangle_{\mathcal{H}} \text{ the inner product in } \mathcal{H} (= \mathcal{H}(k)); \\ &\langle \xi, \eta \rangle_{l^2} = \sum_{x \in V} \xi(x) \eta(x); \\ &\|\xi\|_{l^2}^2 = \sum_{x \in V} |\xi(x)|^2 < \infty; \\ &\|k_x\|_{\mathcal{H}}^2 = k(x, x), \forall x \in V. \end{aligned}$$

To check (4.3), it is enough to prove that

$$\langle A\delta_x, k_y \rangle_{\mathcal{H}} = \langle \delta_x, Bk_y \rangle_{l^2}, \quad \forall x, y \in V. \quad (4.4)$$

Case 1.

$$\text{LHS}_{(4.4)} = \langle \delta_x, k_y \rangle_{\mathcal{H}} = \delta_{xy} = \langle \delta_x, \delta_y \rangle_{l^2} = \text{RHS}_{(4.4)}.$$

Case 2.

$$\text{LHS}_{(4.4)} = \langle \delta_x, k_y \rangle_{\mathcal{H}} = \delta_x(y) - \delta_x(o) = \text{RHS}_{(4.4)}.$$

See (4.1)-(4.2) in Definition 4.1. \square

Notation (closure). Below we shall use the following terminology for the closure of linear operators $\mathcal{H}_1 \xrightarrow{T} \mathcal{H}_2$ where T has dense domain in \mathcal{H}_1 ; and $\text{dom}(T^*)$ is assumed dense in \mathcal{H}_2 . The graph of T is

$$G(T) = \left\{ \begin{pmatrix} h \\ Th \end{pmatrix} \mid h \in \text{dom}(T) \right\} \subseteq \begin{pmatrix} \mathcal{H}_1 \\ \oplus \\ \mathcal{H}_2 \end{pmatrix}.$$

The $\begin{pmatrix} \mathcal{H}_1 \\ \oplus \\ \mathcal{H}_2 \end{pmatrix}$ -closure of $G(T)$ is then the graph of the “closure” of T , written \overline{T} ; in short, $\overline{G(T)} = G(\overline{T})$. We also have $T^{**} = \overline{T}$.

Corollary 4.3. *Let the operators (A, B) be as above.*

- (1) *We have $A \subset B^*$, and $B \subset A^*$; see Fig 4.2.*
- (2) *Moreover, both operators below have dense domains, and are selfadjoint:*
 $\text{dom}(A^*\overline{A})$ *is dense in* l^2 , *and* $A^*\overline{A}$ *is s.a. in* l^2 ;
 $\text{dom}(B^*\overline{B})$ *is dense in* \mathcal{H} , *and* $B^*\overline{B}$ *is s.a. in* \mathcal{H} .
- (3) *Using polar decomposition, we then get:*

$$\overline{A} = V (A^*\overline{A})^{\frac{1}{2}} = (\overline{A}A^*)^{\frac{1}{2}} V \quad (4.5)$$

$$\overline{B} = W (B^*\overline{B})^{\frac{1}{2}} = (\overline{B}B^*)^{\frac{1}{2}} W \quad (4.6)$$

with partial isometries $V : l^2 \longrightarrow \mathcal{H}$, $W : \mathcal{H} \longrightarrow l^2$, and

$$V^*V = I_{l^2} - \text{Projker}(\overline{A}), \text{ and} \quad (4.7)$$

$$W^*W = I_{\mathcal{H}} - \text{Projker}(\overline{B}). \quad (4.8)$$

Proof. The conclusions here follow from the fundamentals regarding the polar decomposition (factorization), in the setting of general unbounded closable operators; see e.g., [Sto51, DS88]. \square

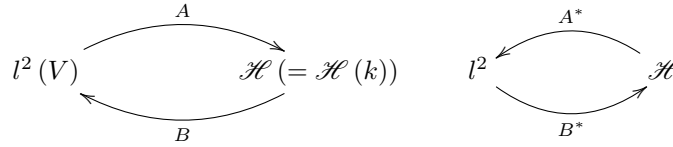


FIGURE 4.2. The symmetric pair (A, B) , with $\mathcal{D}(A) = \text{span}\{\delta_x\}$, and $\mathcal{D}(B) = \text{span}\{k_x\}$.

Remark 4.4. Since $A^*\overline{A}$ is selfadjoint in l^2 , it has a projection valued measure $P^{(A)}$. The following property of $P^{(A)}$ shall be used below: If ψ is a Borel function on $[0, \infty)$, then the functional calculus operator $\psi(A^*\overline{A})$ has the following representation

$$\psi(A^*\overline{A}) = \int_0^\infty \psi(\lambda) P^{(A)}(d\lambda).$$

Given $\xi \in l^2$, we therefore have:

$$\xi \in \text{dom}(\psi(A^*\overline{A})) \iff \int_0^\infty |\psi(\lambda)|^2 \|P^{(A)}(d\lambda)\xi\|_{l^2}^2 < \infty;$$

and in this case,

$$\|\psi(A^*\overline{A})\xi\|_{l^2}^2 = \int_0^\infty |\psi(\lambda)|^2 \|P^{(A)}(d\lambda)\xi\|_{l^2}^2.$$

Remark 4.5. Let (V, k, \mathcal{H}) be as in Definition 4.1 and Theorem 4.8; i.e., we are assuming that $\delta_x \in \mathcal{H}$, $\forall x \in V$. Let (A, B) be the associated symmetric pair; see Proposition 4.2. Set

$$H_{AR}(\mathcal{H}, k) := \{h \in \mathcal{H} \mid (A^*h)(x) = 0, \forall x \in V\} = \ker(A^*), \quad (4.9)$$

and let

$$DEL(\mathcal{H}, k) := \text{the } \mathcal{H}\text{-closure of } \text{span}\{\delta_x \mid x \in V\}. \quad (4.10)$$

Lemma 4.6. *We have the following orthogonal splitting:*

$$\mathcal{H} = H_{AR}(\mathcal{H}, k) \oplus DEL(\mathcal{H}, k) \quad (4.11)$$

Proof. Since $DEL(\mathcal{H}, k) = \{\delta_x \mid x \in V\}^{\perp\perp}$ where “ \perp ” refers to the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, we need only to show that

$$\langle h, \delta_x \rangle_{\mathcal{H}} = 0, \forall x \in V \iff (A^*h)(x) = 0, \forall x \in V. \quad (4.12)$$

But the last eq. (4.12) follows from the duality in Proposition 4.2:

If $\langle h, \delta_x \rangle_{\mathcal{H}} = 0, \forall x \in V$; then $h \in \text{dom}(A^*)$, and

$$(A^*h)(x) = \langle \delta_x, A^*h \rangle_{l^2} = \langle A\delta_x, h \rangle_{\mathcal{H}} = \langle \delta_x, h \rangle_{\mathcal{H}}.$$

The desired conclusion (4.11) is now immediate from this. \square

Question 4.7. *Assume the point-mass property (Def. 2.5), i.e., $\delta_x \in \mathcal{H}, \forall x \in V$. How do we compute the following two positive definite (p.d.) kernels? The p.d. kernel k , with*

$$V \times V \ni (x, y) \longrightarrow k(x, y);$$

and the induced p.d. kernel

$$V \times V \ni (x, y) \longrightarrow \langle \delta_x, \delta_y \rangle_{\mathcal{H}}.$$

Note that we are not assuming that $\delta_x \in \text{dom}(\overline{B})$. How to compute the kernels $k(\cdot, \cdot)$ and $\langle \delta_x, \delta_y \rangle_{\mathcal{H}}$? See details below.

Theorem 4.8. *Let $k : V \times V \longrightarrow \mathbb{C}$ (or \mathbb{R}) be given positive definite, and assume*

$$\delta_x \in \mathcal{H} (= \mathcal{H}(k)), \quad \forall x \in V, \quad (4.13)$$

i.e., \mathcal{H} is a RKHS with point masses. Let (A, B) be the canonical symmetric pair; see Definition 4.1. Then

$$k_x \in \text{dom}(B^*\overline{B}), \quad \forall x \in V. \quad (4.14)$$

Proof. Since $\overline{B}k_x = Bk_x = \delta_x$, the desired conclusion follows if we show that $\delta_x \in \text{dom}(B^*)$, i.e., $\forall x \in V, \exists C_x < \infty$ s.t.

$$|\langle \delta_x, Bf \rangle_{l^2}|^2 \leq C_x \|f\|_{\mathcal{H}}^2, \quad \forall f \in \mathcal{D}(B). \quad (4.15)$$

Fix $x = x_0 \in V$. Now set $f = \sum_y \xi_y k_y \in \mathcal{D}(B)$ (finite sum), then

$$\langle \delta_{x_0}, Bf \rangle_{l^2} = \left\langle \delta_{x_0}, \sum_y \xi_y \delta_y \right\rangle_{l^2} = \xi_{x_0}.$$

But by (4.13) and Lemma 2.4, we get the desired constant $C_x < \infty$ s.t.

$$|\xi_{x_0}|^2 \leq C_{x_0} \sum_{(y,z) \in V \times V} \xi_y \bar{\xi}_z k(y, z) = C_{x_0} \|f\|_{\mathcal{H}}^2;$$

and eq. (4.15) follows, see also (2.12) in Theorem 2.10. \square

Corollary 4.9. *Let V, k, \mathcal{H} be as above, i.e., assuming $\delta_x \in \mathcal{H}, \forall x \in V$.*

- (1) Then $AB : \mathcal{H} \rightarrow \mathcal{H}$ is a symmetric operator in \mathcal{H} with dense domain $\text{span}\{k_x\} = \mathcal{D}(B)$, and $ABk_x = \delta_x \in \mathcal{H}$.
- (2) Moreover, $B^*\bar{B}$ is a selfadjoint extension of AB (as an operator in \mathcal{H} .)

Proof. We have

$$k_x \xrightarrow[\mathcal{H}]{B} \delta_x \xrightarrow[l^2]{A} \delta_x$$

and we proved that $\delta_x \in \text{dom}(B^*)$ in Theorem 4.8, so we conclude that $B^*\bar{B}$ extends AB . By Corollary 4.3, we know that $B^*\bar{B}$ is a selfadjoint operator in \mathcal{H} . \square

Theorem 4.10. Let V, k , and \mathcal{H} be as above, and assume $\delta_x \in \mathcal{H}$, $\forall x \in V$. Then

$$\delta_x \in \text{dom}(A^*\bar{A}), \quad \forall x \in V \quad (4.16)$$

$$\Updownarrow$$

$$\langle \delta_x, \delta_x \rangle_{\mathcal{H}} \in l^2, \quad \forall x \in V. \quad (4.17)$$

Proof. Note (4.17) means

$$\sum_y |\langle \delta_x, \delta_y \rangle_{\mathcal{H}}|^2 < \infty, \quad \forall x \in V. \quad (4.18)$$

Since $\bar{A}\delta_{x_0} = \delta_{x_0}$, we now show that (4.16) holds \iff (4.18) is satisfied. That is, $\exists C_{x_0} < \infty$, s.t.

$$|\langle A\xi, \delta_{x_0} \rangle_{\mathcal{H}}|^2 \leq C_{x_0} \|\xi\|_{l^2}^2 = C_{x_0} \sum_y |\xi_y|^2, \quad \forall (\xi_y) \text{ finitely indexed.} \quad (4.19)$$

But $\text{LHS}_{(4.19)} = \left| \sum_y \bar{\xi}_y \langle \delta_y, \delta_{x_0} \rangle_{\mathcal{H}} \right|^2$, so (4.19) holds $\iff [y \rightarrow \langle \delta_y, \delta_{x_0} \rangle_{\mathcal{H}}] \in l^2$ which is the desired conclusion in (4.17). \square

Remark 4.11. Note that (4.17) is not automatic. Examples showing this? See Lemma 4.12 and Examples 4.21, 6.9 below.

Lemma 4.12. Let V, k, \mathcal{H} be as above, assuming $\delta_x \in \mathcal{H}$, for all $x \in V$. Then

$$\langle \delta_x, \delta_y \rangle_{\mathcal{H}} = \lim_F (K_F^{-1})_{xy},$$

where the RHS is the inductive limit over the filter of all finite subsets F of V , and

$$K_F = (\langle k_x, k_y \rangle_{\mathcal{H}})_{(x,y) \in F \times F} = (k(x, y))_{(x,y) \in F \times F},$$

i.e., the Gramian matrix.

Proof. We have

$$\begin{aligned} \langle \delta_x, \delta_y \rangle_{\mathcal{H}} &= \lim_F \langle \delta_x, P_F(\delta_y) \rangle_{\mathcal{H}} \\ &= \lim_F \left\langle \delta_x, \sum_s (K_F^{-1})_{sy} k_s \right\rangle_{\mathcal{H}} \quad (\text{Lemma 2.6}) \\ &= \lim_F (K_F^{-1})_{xy} \end{aligned}$$

which is the assertion. \square

Infinite square matrices.

Lemma 4.13. Let k, V and $\mathcal{H} (= \mathcal{H}(k))$ be as above, and assume $\delta_x \in \mathcal{H}$ for all $x \in V$. Consider three $\infty \times \infty$ matrices, D , K , and C as follows:

$$D_{xy} := \langle \delta_x, \delta_y \rangle_{\mathcal{H}} \quad (4.20)$$

Set C s.t.

$$\delta_x = \sum_y C_{xy} k_y, \quad (4.21)$$

and let

$$K_{xy} = k(x, y). \quad (4.22)$$

Then,

$$D = CKC^{tr}, \quad (4.23)$$

or equivalently,

$$D_{xx'} = (CKC^{tr})_{xx'}, \quad \forall (x, x') \in V \times V.$$

Moreover, let $I = (\delta_{xy})_{(x,y) \in V \times V}$ be the $\infty \times \infty$ identity matrix, then we have

$$I = CK.$$

Proof. In our discussion of infinite matrices below, for the infinite summations involved, we are making use of the limit considerations which we made precise in the proof of Lemma 4.12 above.

Apply k_z to both sides in (4.21), then

$$\delta_{xz} = \sum_y C_{xy} K_{yz} \iff I = CK. \text{ (matrix product)}$$

Using $\langle \delta_x, k_z \rangle_{\mathcal{H}} = \delta_{xz}$, we get

$$(CKC^{tr})_{xx'} = \sum_y \sum_{y'} C_{xy} K_{yy'} C_{x'y'}. \quad (4.24)$$

Also,

$$\begin{aligned} D_{xx'} & \stackrel{\text{by (4.20)}}{=} \langle \delta_x, \delta_{x'} \rangle_{\mathcal{H}} \\ & \stackrel{\text{by (4.21)}}{=} \left\langle \sum_y C_{xy} k_y, \sum_{y'} C_{x'y'} k_{y'} \right\rangle_{\mathcal{H}} \\ & = \sum_y \sum_{y'} C_{xy} C_{x'y'} \langle k_y, k_{y'} \rangle_{\mathcal{H}} \\ & \stackrel{\text{by (4.22)}}{=} (CKC^{tr})_{xx'} \end{aligned}$$

which is the desired conclusion, see (4.24). \square

In the discussion below we shall consider matrix algebra for “infinite square matrices.” More precisely, we shall apply matrix algebra to pairs of matrices where in each matrix factor, both rows and columns are indexed by the same given countable infinite set V . Nonetheless, matrix multiplication in this context will require our use of the limit considerations from the proof of Lemma 4.12 above. In other words, the infinite sums entail a limit over filters of finite subsets of V , as discussed in the proof of Lemma 4.12.

Lemma 4.14. We have $C = K^{-1}$, or equivalently,

$$(KC)_{xy} = \delta_{xz} \iff \sum_z K_{xz} C_{zy} = \delta_{xy}.$$

Proof. Apply $k_{x'}$ to both sides in (4.21), and we get

$$\delta_{xx'} = \sum_y C_{xy} K_{yx'} = (CK)_{xx'}.$$

□

Corollary 4.15. *Assuming now that the given positive definite function k is real valued, we then get the following:*

$$D = C^{tr} = (K^{tr})^{-1} = K^{-1}.$$

Proof. Note that $K = K^{tr}$ if $k : V \times V \rightarrow \mathbb{R}$ is real valued since

$$k(x, y) = \langle k_x, k_y \rangle = \langle k_y, k_x \rangle = k(y, x);$$

and so $D = (K^{tr})^{-1} = K^{-1}$. □

Spectral Theory. Let V, k , and $\mathcal{H} = \mathcal{H}(k)$ be as above. Let (A, B) be the associated dual pair of operators from Corollary 4.3. Since $A^* \bar{A}$ is selfadjoint in $l^2 = l^2(V)$ with dense domain, it has a canonical l^2 -projection valued measure $P^{(A)}(\cdot)$ defined on the Borel σ -algebra \mathcal{B}_+ of subsets of $[0, \infty)$. We set

$$d\mu_x^{(A)}(\lambda) := \left\| P^{(A)}(d\lambda) \delta_x \right\|_{l^2}^2. \quad (4.25)$$

Lemma 4.16.

(1) *The following conclusions hold for the measure $\mu_x^{(A)}$:*

(a) *Moments of order 0, 1, and 2:*

$$\mu_x^{(A)}([0, \infty)) = 1, \quad \forall x \in V; \quad (4.26)$$

$$\int_0^\infty \lambda d\mu_x^{(A)}(\lambda) = \|\delta_x\|_{\mathcal{H}}^2; \text{ and} \quad (4.27)$$

$$\int_0^\infty \lambda^2 d\mu_x^{(A)}(\lambda) = \sum_{y \in V} |\langle \delta_y, \delta_x \rangle_{\mathcal{H}}|^2 = \|A^* \delta_x\|_{l^2}^2; \quad (4.28)$$

(b) *The covariance of the measure $\mu_x^{(A)}$ is:*

$$\text{cov}(\mu_x^{(A)}) = \|A^* \delta_x\|_{l^2}^2 - \|\delta_x\|_{\mathcal{H}}^4. \quad (4.29)$$

(2) *Moreover, the first moment in (4.27) is finite iff $\delta_x \in \mathcal{H}$. The second moment in (4.28) is finite iff $\delta_x \in \text{dom}(A^*)$.*

Proof.

For (4.26), we have:

$$\mu_x^{(A)}([0, \infty)) = \left\| P^{(A)}([0, \infty)) \delta_x \right\|_{l^2}^2 = \|\delta_x\|_{l^2}^2 = 1.$$

For (4.27), we have:

$$\begin{aligned} \int_0^\infty \lambda d\mu_x^{(A)}(\lambda) & \stackrel{\text{(by Cor. 4.3)}}{=} \left\| (A^* \bar{A})^{\frac{1}{2}} \delta_x \right\|_{l^2}^2 \\ & \stackrel{\text{(by (4.5))}}{=} \left\| V(A^* \bar{A})^{\frac{1}{2}} \delta_x \right\|_{\mathcal{H}}^2 \\ & \stackrel{\text{(by (4.5))}}{=} \|\bar{A} \delta_x\|_{\mathcal{H}}^2 \end{aligned}$$

$$\stackrel{=}{\text{(by Def. 4.1)}} \|\delta_x\|_{\mathcal{H}}^2.$$

For (4.28), if $\delta_x \in \text{dom}(A^*)$, then $A^*\delta_x \in l^2$. Therefore

$$A^*\delta_x = \sum_{y \in V} \langle \delta_y, A^*\delta_x \rangle_{l^2} \delta_y, \quad (l^2\text{-convergence})$$

and

$$\|A^*\delta_x\|_{l^2}^2 = \sum_{y \in V} |\langle \delta_y, A^*\delta_x \rangle_{l^2}|^2.$$

But

$$\langle \delta_y, A^*\delta_x \rangle_{l^2} = \langle A\delta_y, \delta_x \rangle_{\mathcal{H}} = \langle \delta_y, \delta_x \rangle_{\mathcal{H}}. \quad (4.30)$$

Therefore, with the use of Remark 4.4, we arrive at the following:

$$\begin{aligned} \int_0^\infty \lambda^2 d\mu_x^{(A)}(\lambda) &\stackrel{=}{\text{(by Cor. 4.3)}} \|A^*\overline{A}\delta_x\|_{l^2}^2 \\ &\stackrel{=}{\text{(by (4.30))}} \sum_{y \in V} |\langle \delta_y, \delta_x \rangle_{\mathcal{H}}|^2. \end{aligned}$$

The remaining conclusions in the lemma are now immediate from this. \square

Corollary 4.17. *If the equivalent conditions in Theorem 4.10 are satisfied, then, for every $x \in V$, there is a finite positive Borel measure μ_x on $[0, \infty)$ such that*

$$\sum_{y \in V} |\langle \delta_y, \delta_x \rangle_{\mathcal{H}}|^2 = \int_0^\infty \lambda^2 d\mu_x(\lambda), \quad \forall x \in V. \quad (4.31)$$

In general, the \mathcal{H} -norm of δ_x is finite iff the first moment of μ_x is finite.

Proof. Let the condition in the corollary hold. We then make use of the selfadjoint operator $A^*\overline{A}$ from Corollary 4.3. We conclude that $\delta_x \in \text{dom}(A^*\overline{A})$, $\forall x \in V$. Let $P^{(A)}$ denote the projection valued measure obtained from the selfadjoint operator $A^*\overline{A}$, i.e.,

$$A^*\overline{A} = \int_0^\infty \lambda P^{(A)}(\lambda) \quad (4.32)$$

holds on the dense domain $\text{dom}(A^*\overline{A})$; hence if $\delta_x \in \text{dom}(A^*\overline{A})^{\frac{1}{2}}$, we get

$$\left\| (A^*\overline{A})^{\frac{1}{2}} \delta_x \right\|_{l^2}^2 = \langle \overline{A}\delta_x, \overline{A}\delta_x \rangle_{\mathcal{H}} = \langle \delta_x, \delta_x \rangle_{\mathcal{H}} = \|\delta_x\|_{\mathcal{H}}^2.$$

Now set

$$\mu_x(\cdot) = \|P(\cdot)\delta_x\|_{l^2}^2, \quad (4.33)$$

and substitute into (4.32). We get

$$\begin{aligned} \|\delta_x\|_{\mathcal{H}}^2 &= \int_0^\infty \lambda \langle \delta_x, P(d\lambda)\delta_x \rangle_{l^2} \\ &= \int_0^\infty \lambda \|P(d\lambda)\delta_x\|_{l^2}^2 \\ &= \int_0^\infty \lambda d\mu_x(\lambda), \end{aligned}$$

which is the remaining conclusion. \square

Corollary 4.18. *Let k, V, \mathcal{H} , and $P^{(A)}(\cdot)$ be as above; i.e., $P^{(A)}$ is the projection valued measure of the selfadjoint operator $A^* \bar{A}$ in l^2 . For $x, y \in V$, set*

$$d\mu_{x,y}^{(A)}(\lambda) = \left\langle \delta_x, P^{(A)}(d\lambda) \delta_y \right\rangle_{l^2}. \quad (4.34)$$

Then

$$\langle \delta_x, \delta_y \rangle_{\mathcal{H}} = \int_0^\infty \lambda d\mu_{x,y}^{(A)}(\lambda). \quad (4.35)$$

Proof. By Lemma 4.16 and Remark 4.4, we have

$$\int_0^\infty \lambda d\mu_{x,y}^{(A)}(\lambda) = \langle \delta_x, A^* \bar{A} \delta_y \rangle_{l^2} = \langle \delta_x, \delta_y \rangle_{\mathcal{H}}.$$

□

Let (k, V, \mathcal{H}) be as in Theorem 4.8 and Corollary 4.18; and let (A, B) be the associated symmetric pair. Let $B^* \bar{B}$ be the selfadjoint operator in \mathcal{H} , introduced in Corollary 4.3; and let $P^{(B)}(\cdot)$ be the corresponding projection valued measure; i.e., $P^{(B)}(S)$ is a projection in \mathcal{H} , $\forall S \in \mathcal{B}_+ =$ the Borel σ -algebra of subsets of $[0, \infty)$. In particular,

$$B^* \bar{B} = \int_0^\infty \lambda P^{(B)}(d\lambda) \quad (4.36)$$

holds on the dense domain $\text{dom}(B^* \bar{B})$ in \mathcal{H} .

Proposition 4.19. *For $x \in V$, set*

$$d\mu_x^{(B)}(\lambda) = \left\| P^{(B)}(d\lambda) k_x \right\|_{\mathcal{H}}^2. \quad (4.37)$$

Then for the moments of order 0, 1, and 2, we have:

$$\mu_x^{(B)}([0, \infty)) = k(x, x); \quad (4.38)$$

$$\int_0^\infty \lambda d\mu_x^{(B)}(\lambda) = 1, \quad \text{and} \quad (4.39)$$

$$\int_0^\infty \lambda^2 d\mu_x^{(B)}(\lambda) = \|\delta_x\|_{\mathcal{H}}^2. \quad (4.40)$$

Proof. We have

$$\text{LHS}_{(4.38)} = \|k_x\|_{\mathcal{H}}^2 = k(x, x).$$

$$\begin{aligned} \text{LHS}_{(4.39)} &= \left\| (B^* \bar{B})^{\frac{1}{2}} k_x \right\|_{\mathcal{H}}^2 \\ &= \left\| W(B^* \bar{B})^{\frac{1}{2}} k_x \right\|_{l^2}^2 \quad (\text{see (4.6) in Cor. 4.3}) \\ &\stackrel{(\text{by (4.6)})}{=} \left\| \bar{B} k_x \right\|_{l^2}^2 \\ &\stackrel{(\text{by (4.1)})}{=} \|\delta_x\|_{l^2}^2 = 1. \end{aligned}$$

Similarly,

$$\begin{aligned} \text{LHS}_{(4.40)} &= \|B^* \bar{B} k_x\|_{\mathcal{H}}^2 \\ &\stackrel{(\text{by (4.1)})}{=} \|B^* \delta_x\|_{\mathcal{H}}^2 = \|A \delta_x\|_{\mathcal{H}}^2 = \|\delta_x\|_{\mathcal{H}}^2. \end{aligned}$$

We have proved the three moments formulas. □

4.1. Application: Moment analysis of networks with given conductance function. Consider a fixed infinite network as specified in Definitions 3.1-3.2. Recall that $c : E \rightarrow \mathbb{R}_+$ is a fixed conductance function, i.e., $c_{xy} = c_{yx}$, and defined for $\forall (xy) \in E$. We write $x \sim y$ iff (Def.) $(xy) \in E$. Set

$$c(x) = \sum_{y \sim x} c_{xy}. \quad (4.41)$$

The sum in (4.41) may be finite, or infinite. Let $x \in V \setminus \{o\}$, where “ o ” is the chosen base-point in the vertex set V .

Theorem 4.20. *Given (V, E, c) connected, and let $x \in V \setminus \{o\}$. Set $\mu_x^{(A)}(\cdot) = \|P^{(A)}(\cdot) \delta_x\|_{l^2}^2$ (see (4.25)).*

(1) *We have*

$$\delta_x \in \mathcal{H}_E \iff c(x) < \infty, \quad (4.42)$$

and in this case

$$\int_0^\infty \lambda d\mu_x^{(A)}(\lambda) = \|\delta_x\|_{\mathcal{H}_E}^2 = c(x). \quad (4.43)$$

(2) *Assume (4.42) holds for all $x \in V \setminus \{o\}$, then*

$$\int_0^\infty \lambda^2 d\mu_x^{(A)}(\lambda) = \left(c(x)^2 + \sum_{y \sim x} c_{xy}^2\right). \quad (4.44)$$

(3) *For the covariance of $\mu_x^{(A)}$, we have:*

$$\text{cov}\left(\mu_x^{(A)}\right) = \sum_{y \sim x} c_{xy}^2, \quad (4.45)$$

and

$$\text{cov}\left(\mu_x^{(A)}\right) \leq \|\delta_x\|_{\mathcal{H}_E}^4. \quad (4.46)$$

Proof. With the use of Lemma 3.3 and Proposition 3.6, we get the following formulas for the \mathcal{H}_E -inner product, see also Definition 3.2:

$$\langle \delta_x, \delta_y \rangle_{\mathcal{H}_E} = \begin{cases} c(x) & \text{if } y = x \\ -c_{xy} & \text{if } y \sim x \\ 0 & \text{if } y \neq x \text{ and } (xy) \notin E. \end{cases} \quad (4.47)$$

The conclusions (4.42) and (4.43) are immediate from this since $\|\delta_x\|_{\mathcal{H}_E}^2 = c(x)$ follows from (4.47).

Conclusion (4.44) in the theorem follows from (4.28) in Lemma 4.16, and (4.47) above. Indeed,

$$\begin{aligned} \int_0^\infty \lambda^2 d\mu_x^{(A)}(\lambda) &= \|A^* \delta_x\|_{l^2}^2 \\ &= \sum_{y \in V} \left| \langle \delta_y, \delta_x \rangle_{\mathcal{H}_E} \right|^2 \\ &= c(x)^2 + \sum_{y \sim x} c_{xy}^2 \quad (\text{by (4.47)}) \end{aligned}$$

which is the desired conclusion.

For the covariance, we have:

$$\begin{aligned} \int_0^\infty |\lambda - c(x)|^2 d\mu_x^{(A)}(\lambda) &= \int_0^\infty \lambda^2 d\mu_x^{(A)}(\lambda) - c(x)^2 \\ &= \sum_{y \sim x} c_{xy}^2; \quad (\text{by (4.44)}) \end{aligned}$$

thus completing the proof of conclusions (1)-(2) in the statement of the Theorem.

Part (3). Since

$$\sum_{y \sim x} c_{xy}^2 \leq \left(\sum_{y \sim x} c_{xy} \right)^2 = c(x)^2$$

we get estimate (4.46) in part (3) from the Theorem. \square

4.2. Discrete sample points for Brownian motion. We interrupt the general considerations with an example for illustration, choices of discrete sample points for standard Brownian motion.

Example 4.21. Consider $V : 0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \cdots$, a discrete subset of \mathbb{R}_+ , and set

$$k(s, t) = s \wedge t = \min(s, t), \quad \forall s, t \in V.$$

Note that k is the covariance kernel (positive definite) of standard Brownian motion, restricted to the set V . Let $\mathcal{H} (= \mathcal{H}(k))$ be the associated RKHS. (See sect. 6.1 for details.)

For each finite subset $F_n = \{x_1, x_2, \dots, x_n\}$ of V , we have

$$K_n = K^{(F_n)} = \begin{bmatrix} x_1 & x_1 & x_1 & \cdots & x_1 \\ x_1 & x_2 & x_2 & \cdots & x_2 \\ x_1 & x_2 & x_3 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} = (x_i \wedge x_j)_{i,j=1}^n.$$

A direct calculation shows that

$$(K_n^{-1}) = \begin{bmatrix} -\frac{x_2}{x_1^2 - x_1 x_2} & \frac{1}{x_1 - x_2} & 0 & 0 & 0 \\ \frac{1}{x_1 - x_2} & \frac{x_3 - x_1}{(x_1 - x_2)(x_2 - x_3)} & \frac{1}{x_2 - x_3} & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \frac{1}{x_{n-1} - x_n} & \frac{x_{n+1} - x_{n-1}}{(x_{n-1} - x_n)(x_n - x_{n+1})} & \frac{1}{x_n - x_{n+1}} \end{bmatrix}.$$

It follows that

$$[x_j \longrightarrow \langle \delta_{x_i}, \delta_{x_j} \rangle] \in l^2, \quad \forall x_i \in V.$$

See Lemma 4.12.

5. SPECTRAL THEORY: A NECESSARY AND SUFFICIENT CONDITION FOR WHEN THE SYMMETRIC PAIR IS MAXIMAL

We showed in Section 4 that, to every reproducing kernel Hilbert space \mathcal{H} having a countable discrete set of sample points of finite \mathcal{H} -norm, there is a canonically associated symmetric pair of operators (A, B) . In the present section we give a practical necessary and sufficient condition for this symmetric pair to be maximal.

Theorem 5.1. *Let V, k, \mathcal{H} be as above, and assume $\delta_x \in \mathcal{H}, \forall x \in V$. Let (A, B) be the associated symmetric pair of operators from Definition 4.1 and Corollary 4.3. Then TFAE:*

- (1) $\overline{A} = B^*$ and $\overline{B} = A^*$;
- (2) The following implication holds:

$$[h \in \text{dom}(A^*), (A^*h)(x) = -h(x), \forall x \in V] \implies h = 0.$$

Proof. It is enough to consider one of the two conditions in (1). We note that $\overline{B} = A^* \iff$

$$\mathcal{G}(A^*) \ominus \mathcal{G}(B) = 0. \quad (5.1)$$

But

$$\begin{aligned} \begin{pmatrix} h \\ A^*h \end{pmatrix} &\in \mathcal{G}(A^*) \ominus \mathcal{G}(B) \\ &\Downarrow \\ \left\langle \begin{pmatrix} k_x \\ \delta_x \end{pmatrix}, \begin{pmatrix} h \\ A^*h \end{pmatrix} \right\rangle_{\oplus} &= 0, \quad \forall x \in V, \end{aligned} \quad (5.2)$$

where $\langle \cdot, \cdot \rangle_{\oplus}$ is the inner product in $\frac{\mathcal{H}}{l^2}$. Now (5.2) \iff

$$\begin{aligned} \langle k_x, h \rangle_{\mathcal{H}} + \langle \delta_x, A^*h \rangle_{l^2} &= 0, \quad \forall x \in V \\ &\Downarrow \\ h(x) + (A^*h)(x) &\equiv 0, \quad \forall x \in V. \end{aligned}$$

The desired conclusion (1) \iff (2) is now immediate. \square

Corollary 5.2. *Let V, k, \mathcal{H} be as above, and assume $\delta_x \in \mathcal{H}, \forall x \in V$. Let (A, B) be the associated dual pair of operators, with $P^{(A)}$ and $P^{(B)}$ the respective projection valued measures. Set $\mu_x^{(A)}$ and $\mu_x^{(B)}$ as in (4.25) and (4.37). Then we have*

$$\int_0^\infty \lambda d\mu_x^{(A)}(\lambda) = \int_0^\infty \lambda^2 d\mu_x^{(B)}(\lambda) \left(= \|\delta_x\|_{\mathcal{H}}^2 \right). \quad (5.3)$$

Moreover, assume (A, B) is maximal, and $\delta_x \in \text{dom}(B^*\overline{B})$; then

$$\int_0^\infty \lambda^2 d\mu_x^{(A)}(\lambda) = \int_0^\infty \lambda \left\| P^{(B)}(\lambda) \delta_x \right\|_{\mathcal{H}}^2. \quad (5.4)$$

Proof. Eq. (5.3) follows from (4.27) and (4.40).

For (5.4), we have

$$\begin{aligned} \int_0^\infty \lambda^2 d\mu_x^{(A)}(\lambda) &\stackrel{\text{(by (4.28))}}{=} \|A^*\delta_x\|_{l^2}^2 = \|\overline{B}\delta_x\|_{l^2}^2 \\ &= \langle \delta_x, B^*\overline{B}\delta_x \rangle_{\mathcal{H}} \\ &= \int_0^\infty \lambda \left\| P^{(B)}(\lambda) \delta_x \right\|_{\mathcal{H}}^2, \end{aligned}$$

which is (5.4). \square

6. SAMPLE POINT-MASSSES IN CONCRETE MODELS

Suppose $V \subset D \subset \mathbb{R}^d$ where V is countable and discrete, but D is open. In this case, we get two kernels: k on $D \times D$, and $k_V := k|_{V \times V}$ on $V \times V$ by restriction. If $x \in V$, then $k_x^{(V)}(\cdot) = k(\cdot, x)$ is a function on V , while $k_x(\cdot) = k(\cdot, x)$ is a function on D .

This means that the corresponding RKHSs are different, \mathcal{H}_V vs \mathcal{H} , where \mathcal{H}_V = a RKHS of functions on V , and \mathcal{H} = a RKHS of functions on D .

Lemma 6.1. \mathcal{H}_V is isometrically contained in \mathcal{H} via $k_x^{(V)} \mapsto k_x, x \in V$.

Proof. If $F \subset V$ is a finite subset, and $\xi = \xi_F$ is a function on F , then

$$\left\| \sum_{x \in F} \xi(x) k_x^{(V)} \right\|_{\mathcal{H}_V} = \left\| \sum_{x \in F} \xi(x) k_x \right\|_{\mathcal{H}}.$$

The desired result follows from this. (See Proposition 6.23 for the case of *point-mass samples*.) \square

Examples. We are concerned with cases of kernels $k : D \times D \rightarrow \mathbb{R}$ with restriction $k_V : V \times V \rightarrow \mathbb{R}$, where V is a countable discrete subset of D . Typically, for $x \in V$, we may have (restriction) $\delta_x|_V \in \mathcal{H}_V$, but $\delta_x \notin \mathcal{H}$; indeed this happens for the kernel k of standard Brownian motion:

$D = \mathbb{R}_+$;

V = an ordered subset $0 < x_1 < x_2 < \dots < x_i < x_{i+1} < \dots, V = \{x_i\}_{i=1}^\infty$.

In this case, we compute \mathcal{H}_V , and we show that $\delta_{x_i}|_V \in \mathcal{H}_V$; while for \mathcal{H}_m = the Cameron-Martin Hilbert space, we have $\delta_{x_i} \notin \mathcal{H}_m$.

Also note that δ_{x_1} has a different meaning with reference to \mathcal{H}_V vs \mathcal{H}_m . In the first case, it is simply $\delta_{x_1}(y) = \begin{cases} 1 & y = x_1 \\ 0 & y \in V \setminus \{x_1\} \end{cases}$. In the second case, δ_{x_1} is a Schwartz distribution. We shall abuse notation, writing δ_x in both cases.

In the following, we will consider restriction to $V \times V$ of a special continuous p.d. kernel k on $\mathbb{R}_+ \times \mathbb{R}_+$. It is $k(s, t) = s \wedge t = \min(s, t)$. Before we restrict, note that the RKHS of this k is the Cameron-Martin Hilbert space of function f on \mathbb{R}_+ with distribution derivative $f' \in L^2(\mathbb{R}_+)$, and

$$\|f\|_{\mathcal{H}}^2 := \int_0^\infty |f'(t)|^2 dt < \infty. \quad (6.1)$$

For details, see below.

Remark 6.2 (Application). The Hilbert space given by $\|\cdot\|_{\mathcal{H}}^2$ in (6.1) is called the Cameron-Martin Hilbert space, and, as noted, it is the RKHS of $k : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R} : k(s, t) := s \wedge t$. Now pick a discrete subset $V \subset \mathbb{R}_+$; then Lemma 6.1 states that the RKHS of the $V \times V$ restricted kernel, $k^{(V)}$ is isometrically embedded into \mathcal{H} , i.e., setting

$$J^{(V)}(k_x^{(V)}) = k_x, \quad \forall x \in V; \quad (6.2)$$

$J^{(V)}$ extends by “closed span” to an isometry $\mathcal{H}_V \xrightarrow{J^{(V)}} \mathcal{H}$. It further follows from the lemma, that the range of $J^{(V)}$ may have infinite co-dimension.

Note that $P_V := J^{(V)}(J^{(V)})^*$ is the projection onto the range of $J^{(V)}$. The ortho-complement is as follow:

$$\mathcal{H} \ominus \mathcal{H}_V = \{\psi \in \mathcal{H} \mid \psi(x) = 0, \forall x \in V\}. \quad (6.3)$$

Example 6.3. Let k and $k^{(V)}$ be as in (6.2), and set $V := \pi\mathbb{Z}_+$, i.e., integer multiples of π . Then easy generators of wavelet functions [BJ02] yield non-zero functions ψ on \mathbb{R}_+ such that

$$\psi \in \mathcal{H} \ominus \mathcal{H}_V. \quad (6.4)$$

More precisely,

$$0 < \int_0^\infty |\psi'(t)|^2 dt < \infty, \quad (6.5)$$

where ψ' is the distribution (weak) derivative; and

$$\psi(n\pi) = 0, \quad \forall n \in \mathbb{Z}_+. \quad (6.6)$$

An explicit solution to (6.4)-(6.6) is

$$\psi(t) = \prod_{n=1}^\infty \cos\left(\frac{t}{2^n}\right) = \frac{\sin t}{t}, \quad \forall t \in \mathbb{R}. \quad (6.7)$$

From this, one easily generates an infinite-dimensional set of solutions.

6.1. Sample points in Brownian motion. Consider the covariance function of standard Brownian motion B_t , $t \in [0, \infty)$, i.e., a Gaussian process $\{B_t\}$ with mean zero and covariance function

$$\mathbb{E}(B_s B_t) = s \wedge t = \min(s, t). \quad (6.8)$$

We now show that the restriction of (6.8) to $V \times V$ for an ordered subset (we fix such a set V):

$$V : 0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \cdots \quad (6.9)$$

has the discrete mass property (Definition 2.5).

Set $\mathcal{H}_V = RKHS(k|_{V \times V})$,

$$k_V(x_i, x_j) = x_i \wedge x_j. \quad (6.10)$$

We consider the set $F_n = \{x_1, x_2, \dots, x_n\}$ of finite subsets of V , and

$$K_n = k^{(F_n)} = \begin{bmatrix} x_1 & x_1 & x_1 & \cdots & x_1 \\ x_1 & x_2 & x_2 & \cdots & x_2 \\ x_1 & x_2 & x_3 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} = (x_i \wedge x_j)_{i,j=1}^n. \quad (6.11)$$

We will show that condition 3 in Theorem 2.10 holds for k_V . For this, we must compute all the determinants, $D_n = \det(K_F)$ etc. ($n = \#F$), see Corollary 2.11.

Lemma 6.4.

$$D_n = \det\left((x_i \wedge x_j)_{i,j=1}^n\right) = x_1(x_2 - x_1)(x_3 - x_2) \cdots (x_n - x_{n-1}). \quad (6.12)$$

Proof. Induction. In fact,

$$\begin{bmatrix} x_1 & x_1 & x_1 & \cdots & x_1 \\ x_1 & x_2 & x_2 & \cdots & x_2 \\ x_1 & x_2 & x_3 & \cdots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} \sim \begin{bmatrix} x_1 & 0 & 0 & \cdots & 0 \\ 0 & x_2 - x_1 & 0 & \cdots & 0 \\ 0 & 0 & x_3 - x_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & x_n - x_{n-1} \end{bmatrix},$$

unitary equivalence in finite dimensions. \square

Lemma 6.5. *Let*

$$\zeta_{(n)} := K_n^{-1}(\delta_{x_1})(\cdot) \quad (6.13)$$

be as in (2.9), so that

$$\|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}_V}^2 = \zeta_{(n)}(x_1). \quad (6.14)$$

Then,

$$\begin{aligned} \zeta_{(1)}(x_1) &= \frac{1}{x_1} \\ \zeta_{(n)}(x_1) &= \frac{x_2}{x_1(x_2 - x_1)}, \quad \text{for } n = 2, 3, \dots, \end{aligned}$$

and

$$\|\delta_{x_1}\|_{\mathcal{H}_V}^2 = \frac{x_2}{x_1(x_2 - x_1)}.$$

Proof. A direct computation shows the $(1, 1)$ minor of the matrix K_n^{-1} is

$$D'_{n-1} = \det \left((x_i \wedge x_j)_{i,j=2}^n \right) = x_2(x_3 - x_2)(x_4 - x_3) \cdots (x_n - x_{n-1}) \quad (6.15)$$

and so

$$\begin{aligned} \zeta_{(1)}(x_1) &= \frac{1}{x_1}, \quad \text{and} \\ \zeta_{(2)}(x_1) &= \frac{x_2}{x_1(x_2 - x_1)} \\ \zeta_{(3)}(x_1) &= \frac{x_2(x_3 - x_2)}{x_1(x_2 - x_1)(x_3 - x_2)} = \frac{x_2}{x_1(x_2 - x_1)} \\ \zeta_{(4)}(x_1) &= \frac{x_2(x_3 - x_2)(x_4 - x_3)}{x_1(x_2 - x_1)(x_3 - x_2)(x_4 - x_3)} = \frac{x_2}{x_1(x_2 - x_1)} \\ &\vdots \end{aligned}$$

The result follows from this, and from Corollary 2.9. \square

Corollary 6.6. $P_{F_n}(\delta_{x_1}) = P_{F_2}(\delta_{x_1}), \forall n \geq 2$. *Therefore,*

$$\delta_{x_1} \in \mathcal{H}_V^{(F_2)} := \text{span}\{k_{x_1}^{(V)}, k_{x_2}^{(V)}\} \quad (6.16)$$

and

$$\delta_{x_1} = \zeta_{(2)}(x_1)k_{x_1}^{(V)} + \zeta_{(2)}(x_2)k_{x_2}^{(V)} \quad (6.17)$$

where

$$\zeta_{(2)}(x_i) = K_2^{-1}(\delta_{x_1})(x_i), \quad i = 1, 2.$$

Specifically,

$$\zeta_{(2)}(x_1) = \frac{x_2}{x_1(x_2 - x_1)} \quad (6.18)$$

$$\zeta_{(2)}(x_2) = \frac{-1}{x_2 - x_1}; \quad (6.19)$$

and

$$\|\delta_{x_1}\|_{\mathcal{H}_V}^2 = \frac{x_2}{x_1(x_2 - x_1)}. \quad (6.20)$$

Proof. Follows from the lemma. Note that

$$\zeta_n(x_1) = \|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2$$

and $\zeta_{(1)}(x_1) \leq \zeta_{(2)}(x_1) \leq \dots$, since $F_n = \{x_1, x_2, \dots, x_n\}$. In particular, $\frac{1}{x_1} \leq \frac{x_2}{x_1(x_2 - x_1)}$, which yields (6.20). \square

Remark 6.7. We showed that $\delta_{x_1} \in \mathcal{H}_V$, $V = \{x_1 < x_2 < \dots\} \subset \mathbb{R}_+$, with the restriction of $s \wedge t =$ the covariance kernel of Brownian motion.

The same argument also shows that $\delta_{x_i} \in \mathcal{H}_V$ when $i > 1$. We only need to modify the index notation from the case of the proof for $\delta_{x_1} \in \mathcal{H}_V$. The details are sketched below.

Fix $V = \{x_i\}_{i=1}^\infty$, $x_1 < x_2 < \dots$, then

$$P_{F_n}(\delta_{x_i}) = \begin{cases} 0 & \text{if } n < i - 1 \\ \sum_{s=1}^n (K_{F_n}^{-1} \delta_{x_i})(x_s) k_{x_s} & \text{if } n \geq i \end{cases}$$

and

$$\|P_{F_n}(\delta_{x_i})\|_{\mathcal{H}}^2 = \begin{cases} 0 & \text{if } n < i - 1 \\ \frac{1}{x_i - x_{i-1}} & \text{if } n = i \\ \frac{x_{i+1} - x_{i-1}}{(x_i - x_{i-1})(x_{i+1} - x_i)} & \text{if } n > i \end{cases}$$

Conclusion.

$$\delta_{x_i} \in \text{span} \left\{ k_{x_{i-1}}^{(V)}, k_{x_i}^{(V)}, k_{x_{i+1}}^{(V)} \right\}, \quad \text{and} \quad (6.21)$$

$$\|\delta_{x_i}\|_{\mathcal{H}}^2 = \frac{x_{i+1} - x_{i-1}}{(x_i - x_{i-1})(x_{i+1} - x_i)}. \quad (6.22)$$

Corollary 6.8. *Let $V \subset \mathbb{R}_+$ be countable. If $x_a \in V$ is an accumulation point (from V), then $\|\delta_a\|_{\mathcal{H}_V} = \infty$.*

Example 6.9. An illustration for $0 < x_1 < x_2 < x_3 < x_4$:

$$\begin{aligned} P_F(\delta_{x_3}) &= \sum_{y \in F} \zeta^{(F)}(y) k_y(\cdot) \\ \zeta^{(F)} &= K_F^{-1} \delta_{x_3}. \end{aligned}$$

That is,

$$\underbrace{\begin{bmatrix} x_1 & x_1 & x_1 & x_1 \\ x_1 & x_2 & x_2 & x_2 \\ x_1 & x_2 & x_3 & x_3 \\ x_1 & x_2 & x_3 & x_4 \end{bmatrix}}_{(K_F(x_i, x_j))_{i,j=1}^4} \begin{bmatrix} \zeta^{(F)}(x_1) \\ \zeta^{(F)}(x_2) \\ \zeta^{(F)}(x_3) \\ \zeta^{(F)}(x_4) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

and

$$\begin{aligned} \zeta^{(F)}(x_3) &= \frac{x_1(x_2 - x_1)(x_4 - x_2)}{x_1(x_2 - x_1)(x_3 - x_2)(x_4 - x_3)} \\ &= \frac{x_4 - x_2}{(x_3 - x_2)(x_4 - x_3)} = \|\delta_{x_3}\|_{\mathcal{H}}^2. \end{aligned}$$

Example 6.10 (Sparse sample-points). Let $V = \{x_i\}_{i=1}^\infty$, where

$$x_i = \frac{i(i-1)}{2}, \quad i \in \mathbb{N}.$$

It follows that $x_{i+1} - x_i = i$, and so

$$\|\delta_{x_i}\|_{\mathcal{H}}^2 = \frac{x_{i+1} - x_{i-1}}{(x_i - x_{i-1})(x_{i+1} - x_i)} = \frac{2i - 1}{(i - 1)i} \xrightarrow{i \rightarrow \infty} 0.$$

We conclude that $\|\delta_{x_i}\|_{\mathcal{H}} \xrightarrow{i \rightarrow \infty} 0$ if the set $V = \{x_i\}_{i=1}^{\infty} \subset \mathbb{R}_+$ is sparse.

Now, some general facts:

Lemma 6.11. *Let $k : V \times V \rightarrow \mathbb{C}$ be p.d., and let \mathcal{H} be the corresponding RKHS. If $x_1 \in V$, and if δ_{x_1} has a representation as follows:*

$$\delta_{x_1} = \sum_{y \in V} \zeta^{(x_1)}(y) k_y, \quad (6.23)$$

then

$$\|\delta_{x_1}\|_{\mathcal{H}}^2 = \zeta^{(x_1)}(x_1). \quad (6.24)$$

Proof. Substitute both sides of (6.23) into $\langle \delta_{x_1}, \cdot \rangle_{\mathcal{H}}$ where $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ denotes the inner product in \mathcal{H} . \square

Example 6.12 (Application). Suppose $V = \cup_n F_n$, $F_n \subset F_{n+1}$, where each $F_n \in \mathcal{F}(V)$, then if $x_1 \in F_n$, we have

$$P_{F_n}(\delta_{x_1}) = \sum_{y \in F_n} \langle x_1, K_{F_n}^{-1} y \rangle_{l^2} k_y \quad (6.25)$$

and

$$\|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2 = \langle x_1, K_{F_n}^{-1} x_1 \rangle_{l^2} = (K_{F_n}^{-1} \delta_{x_1})(x_1) \quad (6.26)$$

and the expression $\|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2$ is monotone in n , i.e.,

$$\|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2 \leq \|P_{F_{n+1}}(\delta_{x_1})\|_{\mathcal{H}}^2 \leq \dots \leq \|\delta_{x_1}\|_{\mathcal{H}}^2$$

with

$$\sup_{n \in \mathbb{N}} \|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2 = \lim_{n \rightarrow \infty} \|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2 = \|\delta_{x_1}\|_{\mathcal{H}}^2.$$

For other applications of reproducing kernel Hilbert spaces to the analysis of Gaussian processes, see e.g., [JS09, JP13].

Question 6.13. *Let $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ be positive definite, and let $V \subset \mathbb{R}^d$ be a countable discrete subset, e.g., $V = \mathbb{Z}^d$. When does $k|_{V \times V}$ have the discrete mass property?*

Examples of the affirmative, or not, will be discussed below.

6.2. Discrete RKHSs from restrictions. Let $D := [0, \infty)$, and $k : D \times D \rightarrow \mathbb{R}$, with

$$k(x, y) = x \wedge y = \min(x, y).$$

Restrict to $V := \{0\} \cup \mathbb{Z}_+ \subset D$, i.e., consider

$$k^{(V)} = k|_{V \times V}.$$

$\mathcal{H}(k)$: Cameron-Martin Hilbert space, consisting of functions $f \in L^2(\mathbb{R})$ s.t.

$$\int_0^\infty |f'(x)|^2 dx < \infty, \quad f(0) = 0.$$

$\mathcal{H}_V := \mathcal{H}(k_V)$. Note that

$$f \in \mathcal{H}(k_V) \iff \sum_n |f(n) - f(n+1)|^2 < \infty.$$

Lemma 6.14. *We have $\delta_n = 2k_n - k_{n+1} - k_{n-1} \in \mathcal{H}_V$.*

Proof. Introduce the discrete Laplacian $\Delta = \Delta_c$ (see (3.5)), i.e.,

$$(\Delta f)(x) = \sum_{y \sim x} c_{xy} (f(x) - f(y)),$$

defined for all functions f on $V = \{0\} \cup \mathbb{Z}_+ \subset D$, and $c : E \rightarrow \mathbb{R}_+$ is the corresponding conductance. Setting $c \equiv 1$, we get

$$(\Delta f)(n) = 2f(n) - f(n-1) - f(n+1).$$

But, by (3.7) in Proposition 3.6, we have $\Delta k_n = \delta_n$, and the assertion of the lemma follows from this. Note that

$$\langle 2k_n - k_{n+1} - k_{n-1}, k_m \rangle_{\mathcal{H}_V} = \langle \delta_n, k_m \rangle_{\mathcal{H}_V} = \delta_{n,m}.$$

□

Remark 6.15. The same argument as in the proof of the lemma shows (*mutatis mutandis*) that any ordered discrete countable infinite subset $V \subset [0, \infty)$ yields

$$\mathcal{H}_V := \mathcal{H}(k|_{V \times V})$$

as a RKHS which is discrete in that (Definition 2.5) if $V = \{x_i\}_{i=1}^\infty$, $x_i \in \mathbb{R}_+$, then $\delta_{x_i} \in \mathcal{H}_V$, $\forall i \in \mathbb{N}$.

Proof. Fix vertices $V = \{x_i\}_{i=1}^\infty$,

$$0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \infty, \quad x_i \rightarrow \infty. \quad (6.27)$$

Assign conductance

$$c_{i,i+1} = c_{i+1,i} = \frac{1}{x_{i+1} - x_i} \left(= \frac{1}{\text{dist}} \right) \quad (6.28)$$

Let

$$\begin{aligned} (\Delta f)(x_i) &= \left(\frac{1}{x_{i+1} - x_i} + \frac{1}{x_i - x_{i-1}} \right) f(x_i) \\ &\quad - \frac{1}{x_i - x_{i-1}} f(x_{i-1}) - \frac{1}{x_{i+1} - x_i} f(x_{i+1}) \end{aligned} \quad (6.29)$$

Equivalently,

$$(\Delta f)(x_i) = (c_{i,i+1} + c_{i,i-1}) f(x_i) - c_{i,i-1} f(x_{i-1}) - c_{i,i+1} f(x_{i+1}). \quad (6.30)$$

Then, with (6.30) we have:

$$\Delta k_{x_i} = \delta_{x_i}$$

where $k(\cdot, \cdot) = \text{restriction of } s \wedge t \text{ from } [0, \infty) \times [0, \infty) \text{ to } V \times V$; and therefore

$$\delta_{x_i} = (c_{i,i+1} + c_{i,i-1}) k_{x_i} - c_{i,i+1} k_{x_{i+1}} - c_{i,i-1} k_{x_{i-1}} \in \mathcal{H}_V \quad (6.31)$$

as the right-side in the last equation is a finite sum. Note that now the RKHS is

$$\mathcal{H}_V = \left\{ f : V \rightarrow \mathbb{C} \mid \sum_{i=1}^\infty c_{i,i+1} |f(x_{i+1}) - f(x_i)|^2 < \infty \right\}.$$

□

6.3. Brownian bridge. Let $D := (0, 1)$ be the open interval $0 < t < 1$, and set

$$k_{bridge}(s, t) := s \wedge t - st; \quad (6.32)$$

then (6.32) is the covariance function for the Brownian bridge $B_{bri}(t)$, i.e.,

$$B_{bri}(0) = B_{bri}(1) = 0 \quad (6.33)$$

$$B_{bri}(t) = (1-t) B\left(\frac{t}{1-t}\right), \quad 0 < t < 1; \quad (6.34)$$

where $B(t)$ is Brownian motion; see Lemma 6.1.

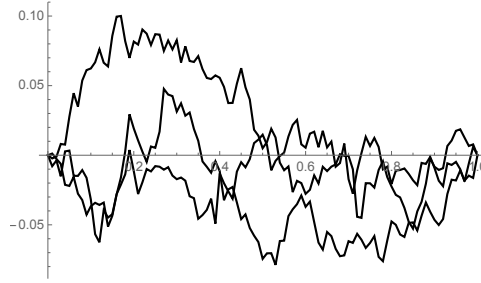


FIGURE 6.1. Brownian bridge $B_{bri}(t)$, a simulation of three sample paths of the Brownian bridge.

The corresponding Cameron-Martin space is now

$$\mathcal{H}_{bri} = \{f \text{ on } [0, 1]; f' \in L^2(0, 1), f(0) = f(1) = 0\} \quad (6.35)$$

with

$$\|f\|_{\mathcal{H}_{bri}}^2 := \int_0^1 |f'(s)|^2 ds < \infty. \quad (6.36)$$

If $V = \{x_i\}_{i=1}^\infty$, $x_1 < x_2 < \dots < 1$, is the discrete subset of D , then we have for $F_n \in \mathcal{F}(V)$, $F_n = \{x_1, x_2, \dots, x_n\}$,

$$K_{F_n} = (k_{bridge}(x_i, x_j))_{i,j=1}^n, \quad (6.37)$$

see (6.32), and

$$\det K_{F_n} = x_1(x_2 - x_1) \cdots (x_n - x_{n-1})(1 - x_n). \quad (6.38)$$

As a result, we get $\delta_{x_i} \in \mathcal{H}_V^{(bri)}$ for all i , and

$$\|\delta_{x_i}\|_{\mathcal{H}_V^{(bri)}}^2 = \frac{x_{i+1} - x_{i-1}}{(x_{i+1} - x_i)(x_i - x_{i-1})}.$$

Note $\lim_{x_i \rightarrow 1} \|\delta_{x_i}\|_{\mathcal{H}_V^{(bri)}}^2 = \infty$.

6.4. Binomial RKHS. It is possible to associate a positive definite kernel (see Definition 6.16) to the standard binomial coefficients. In this section we outline the properties of this kernel and its reproducing kernel Hilbert space. Among the conclusions is that in this RKHS, the point-masses have infinite \mathcal{H} -norm.

Definition 6.16. Let $V = \mathbb{Z}_+ \cup \{0\}$; and

$$k_b(x, y) := \sum_{n=0}^{x \wedge y} \binom{x}{n} \binom{y}{n}, \quad (x, y) \in V \times V.$$

where $\binom{x}{n} = \frac{x(x-1)\cdots(x-n+1)}{n!}$ denotes the standard binomial coefficient from the binomial expansion.

Let $\mathcal{H} = \mathcal{H}(k_b)$ be the corresponding RKHS. Set

$$e_n(x) = \begin{cases} \binom{x}{n} & \text{if } n \leq x \\ 0 & \text{if } n > x. \end{cases} \quad (6.39)$$

Lemma 6.17 ([AJ15]).

- (i) $e_n(\cdot) \in \mathcal{H}$, $n \in V$;
- (ii) $\{e_n\}_{n \in V}$ is an orthonormal basis (ONB) in the Hilbert space \mathcal{H} .
- (iii) Set $F_n = \{0, 1, 2, \dots, n\}$, and

$$P_{F_n} = \sum_{k=0}^n |e_k\rangle\langle e_k| \quad (6.40)$$

or equivalently

$$P_{F_n} f = \sum_{k=0}^n \langle e_k, f \rangle_{\mathcal{H}} e_k. \quad (6.41)$$

then,

(iv) Formula (6.41) is well defined for all functions $f : V \rightarrow \mathbb{C}$, $f \in \mathcal{F}unc(V)$.

(v) Given $f \in \mathcal{F}unc(V)$; then

$$f \in \mathcal{H} \iff \sum_{k=0}^{\infty} |\langle e_k, f \rangle_{\mathcal{H}}|^2 < \infty; \quad (6.42)$$

and, in this case,

$$\|f\|_{\mathcal{H}}^2 = \sum_{k=0}^{\infty} |\langle e_k, f \rangle_{\mathcal{H}}|^2.$$

Fix $x_1 \in V$, then we shall apply Lemma 6.17 to the function $f_1 = \delta_{x_1}$ (in $\mathcal{F}unc(V)$), $f_1(y) = \begin{cases} 1 & \text{if } y = x_1 \\ 0 & \text{if } y \neq x_1. \end{cases}$

Theorem 6.18. We have

$$\|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2 = \sum_{k=x_1}^n \binom{k}{x_1}^2.$$

The proof of the theorem will be subdivided in steps; see below.

Lemma 6.19 ([AJ15]).

(i) For $\forall m, n \in V$, such that $m \leq n$, we have

$$\delta_{m,n} = \sum_{j=m}^n (-1)^{m+j} \binom{n}{j} \binom{j}{m}. \quad (6.43)$$

(ii) For all $n \in \mathbb{Z}_+$, the inverse of the following lower triangle matrix is this:
With (see Figure 6.2)

$$L_{xy}^{(n)} = \begin{cases} \binom{x}{y} & \text{if } y \leq x \leq n \\ 0 & \text{if } x < y \end{cases} \quad (6.44)$$

we have:

$$\left(L^{(n)}\right)_{xy}^{-1} = \begin{cases} (-1)^{x-y} \binom{x}{y} & \text{if } y \leq x \leq n \\ 0 & \text{if } x < y. \end{cases} \quad (6.45)$$

Notation: The numbers in (6.45) are the entries of the matrix $(L^{(n)})^{-1}$.

Proof. In rough outline, (ii) follows from (i). \square

Corollary 6.20. Let k_b , \mathcal{H} , and $n \in \mathbb{Z}_+$ be as above with the lower triangle matrix L_n . Set

$$K_n(x, y) = k_b(x, y), \quad (x, y) \in F_n \times F_n, \quad (6.46)$$

i.e., an $(n+1) \times (n+1)$ matrix.

(i) Then K_n is invertible with

$$K_n^{-1} = (L_n^{tr})^{-1} (L_n)^{-1}; \quad (6.47)$$

an (upper triangle) \times (lower triangle) factorization.

(ii) For the diagonal entries in the $(n+1) \times (n+1)$ matrix K_n^{-1} , we have:

$$\langle x, K_n^{-1} x \rangle_{l^2} = \sum_{k=x}^n \binom{k}{x}^2$$

Conclusion: Since

$$\|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2 = \langle x_1, K_n^{-1} x_1 \rangle_{\mathcal{H}} \quad (6.48)$$

for all $x_1 \in F_n$, we get

$$\begin{aligned} \|P_{F_n}(\delta_{x_1})\|_{\mathcal{H}}^2 &= \sum_{k=x_1}^n \binom{k}{x_1}^2 \\ &= 1 + \binom{x_1+1}{x_1}^2 + \binom{x_1+2}{x_1}^2 + \cdots + \binom{n}{x_1}^2; \end{aligned} \quad (6.49)$$

and therefore,

$$\|\delta_{x_1}\|_{\mathcal{H}}^2 = \sum_{k=x_1}^{\infty} \binom{k}{x_1}^2 = \infty.$$

In other words, no δ_x is in \mathcal{H} .

$$L^{(n)} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & 0 & & & & \vdots & \vdots & \vdots \\ 1 & 3 & 3 & 1 & \ddots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & 1 & 0 & & \vdots & \vdots \\ 1 & \cdots & \binom{x}{y} & \binom{x}{y+1} & \cdots & * & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \ddots & 0 & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & & 1 & 0 \\ 1 & \cdots & \binom{n}{y} & \binom{n}{y+1} & \cdots & \cdots & \cdots & \cdots & n & 1 \end{bmatrix}$$

FIGURE 6.2. The matrix L_n is simply a truncated Pascal triangle, arranged to fit into a lower triangular matrix.

6.5. Classical RKHSs with point-mass samples and interpolation.

Definition 6.21. Let \mathcal{H} be a RKHS (or a relative RKHS) defined from a positive definite kernel $k(x, y)$, $(x, y) \in V \times V$. A (discrete) subset $S \subset V$ is said to be a set of *point-mass samples* iff the span of $\{k_x \mid x \in S\}$ is dense in \mathcal{H} .

Lemma 6.22. Let V, k and \mathcal{H} be as stated in Definition 6.21, and assume $S \subset V$ is a countable discrete subset; then S is a point-mass sample set if $\exists \epsilon \in \mathbb{R}_+$ such that

$$\sum_{s \in S} |f(s)|^2 \geq \epsilon \|f\|_{\mathcal{H}}^2, \quad \forall f \in \mathcal{H}. \quad (6.50)$$

Proof. To show that $\text{span}\{k_s \mid s \in S\}$ is dense in \mathcal{H} , we need only verify that if

$$0 = f(s) = \langle k_s, f \rangle_{\mathcal{H}}, \quad \forall s \in S,$$

then $f \equiv 0$ in \mathcal{H} . But this conclusion is immediate from the estimate (6.50). \square

Proposition 6.23. Let $k : V \times V \rightarrow \mathbb{C}$ be positive definite, and let \mathcal{H} be the corresponding RKHS. Let $S \subset V$ be a set of point-mass samples. Then the following holds for the restricted kernel function $k^{(S)}$, defined by

$$k^{(S)}(s, t) := k(s, t), \quad \forall (s, t) \in S \times S : \quad (6.51)$$

If $\mathcal{H}^{(S)}$ denotes the RKHS of $k^{(S)}$, then the assignment

$$W^{(S)} k_s^{(S)} := k_s, \quad s \in S \quad (6.52)$$

extends by linearity and norm-closure to an isometry $W^{(S)}$ of $\mathcal{H}^{(S)}$ onto \mathcal{H} .

Proof. For finite subsets $F \subset S$, and $\{\xi_s\}_{s \in F}$, we have the following:

$$\left\| \sum_{s \in F} \xi_s k_s^{(S)} \right\|_{\mathcal{H}^{(S)}} = \left\| \sum_{s \in F} \xi_s k_s \right\|_{\mathcal{H}}. \quad (6.53)$$

Hence $W^{(S)}$ in (6.52) extends by linearity and closure to an isometry $W^{(S)} : \mathcal{H}^{(S)} \rightarrow \mathcal{H}$. (Also see Lemma 6.1.)

Since the range $\text{ran}(W^{(S)}) = \{W^{(S)}h^{(S)} \mid h^{(S)} \in \mathcal{H}^{(S)}\}$ is automatically closed in \mathcal{H} , we need only prove that $\text{ran}(W^{(S)})$ is dense in \mathcal{H} ; i.e., $\mathcal{H} \ominus \text{ran}(W^{(S)}) = 0$. By (6.51)-(6.52), we must prove that, if $f \in \mathcal{H}$, and

$$f(s) = \langle k_s, f \rangle_{\mathcal{H}} = 0, \quad \forall s \in S,$$

then $f = 0$ in \mathcal{H} . But the last conclusion is immediate from the condition on the set S from Definition 6.21. \square

Proposition 6.24 (Interpolation). *Let (V, k, \mathcal{H}) be as above, and let $S \subset V$ be a sample set, i.e., satisfying the condition in Definition 6.21. Let (A, B) be the associated dual pair of operators; see Lemma 4.6.*

Then the following interpolation formula holds for $f \in \mathcal{H}$:

$$f = \sum_{s \in S} (A^* f)(s) k_s, \quad (6.54)$$

or equivalently,

$$f(x) = \sum_{s \in S} (A^* f)(s) k(x, s), \quad \forall x \in V; \quad (6.55)$$

and convergence in (6.54)-(6.55) holds iff

$$\sum_{(s,t) \in S \times S} \overline{(A^* f)(s)} (A^* f)(t) k(s, t) < \infty. \quad (6.56)$$

When (6.56) holds, then $\|f\|_{\mathcal{H}}^2 = \text{LHS}_{(6.56)}$.

Proof. Suppose $f = \sum_{s \in S} C_s k_s$ ($\in \mathcal{H}$) is a finite sum-representation; then the coefficients $\{C_s\}_{s \in S}$ are unique. Indeed, if $t \in S$, then

$$\langle \delta_t, f \rangle_{\mathcal{H}} = \sum_{s \in S} C_s \underbrace{\langle \delta_t, k_s \rangle_{\mathcal{H}}}_{\delta_{t,s}} = C_t; \text{ and}$$

$$\langle \delta_t, f \rangle_{\mathcal{H}} = \langle A \delta_t, f \rangle_{\mathcal{H}} = \langle \delta_t, A^* f \rangle_{l^2} = (A^* f)(t), \quad t \in S.$$

\square

The next example shows that there are many RKHSs (k, \mathcal{H}, V) which satisfy the condition in Definition 6.21 for a variety of countably discrete sample sets $S \subset V$; but nonetheless, the point-masses δ_x are not in \mathcal{H} , i.e., $\|\delta_x\|_{\mathcal{H}} = \infty$ for all $x \in V$.

Example 6.25. Let $V = \mathbb{R}$, and let $\mathcal{H} = \{f \in L^2(\mathbb{R}) \mid \text{suppt} \hat{f} \subset [-\frac{1}{2}, \frac{1}{2}]\}$, where \hat{f} denotes the Fourier transform. This RKHS is said to be a band-limited Hilbert space. It is known that then

$$k(x, y) = \frac{\sin \pi(x - y)}{\pi(x - y)}, \quad x, y \in \mathbb{R} \quad (6.57)$$

is a positive definite kernel turning \mathcal{H} into a RKHS.

Moreover, $\{k_n \mid n \in \mathbb{Z}\}$ is then a set of point-mass samples. In fact, $\{k_n\}_{n \in \mathbb{Z}}$ is an orthonormal basis (ONB) in \mathcal{H} , and

$$f(x) = \sum_{n \in \mathbb{Z}} k(n, x) f(n), \quad \forall f \in \mathcal{H} \quad (6.58)$$

holds. Note that (6.58) is Shannon’s sampling formula, and we have

$$\|f\|_{\mathcal{H}}^2 = \sum_{n \in \mathbb{Z}} |f(n)|^2, \quad \forall f \in \mathcal{H}.$$

It is also known that in addition to \mathbb{Z} , there are many other choices of discrete point-mass samples for \mathcal{H} .

For related, recent studies of sampling spaces corresponding to irregular distribution of sample-points, see e.g., [JS13, JS12].

Acknowledgement. The co-authors thank the following colleagues for helpful and enlightening discussions: Professors Sergii Bezuglyi, Ilwoo Cho, Paul Muhly, Myung-Sin Song, Wayne Polyzou, and members in the Math Physics seminar at The University of Iowa.

REFERENCES

- [ABDdS93] Daniel Alpay, Vladimir Bolotnikov, Aad Dijkstra, and Henk de Snoo, *On some operator colligations and associated reproducing kernel Hilbert spaces*, Operator extensions, interpolation of functions and related topics, Oper. Theory Adv. Appl., vol. 61, Birkhäuser, Basel, 1993, pp. 1–27. MR 1246577 (94i:47018)
- [AD92] Daniel Alpay and Harry Dym, *On reproducing kernel spaces, the Schur algorithm, and interpolation in a general class of domains*, Operator theory and complex analysis (Sapporo, 1991), Oper. Theory Adv. Appl., vol. 59, Birkhäuser, Basel, 1992, pp. 30–77. MR 1246809 (94j:46034)
- [AD93] ———, *On a new class of structured reproducing kernel spaces*, J. Funct. Anal. **111** (1993), no. 1, 1–28. MR 1200633 (94g:46035)
- [AJ15] Daniel Alpay and Palle Jorgensen, *Reproducing kernel Hilbert spaces generated by the binomial coefficients*. To appear., ArXiv e-prints (2015).
- [AJSV13] Daniel Alpay, Palle Jorgensen, Ron Seager, and Dan Volok, *On discrete analytic functions: products, rational functions and reproducing kernels*, J. Appl. Math. Comput. **41** (2013), no. 1-2, 393–426. MR 3017129
- [AJV14] Daniel Alpay, Palle Jorgensen, and Dan Volok, *Relative reproducing kernel Hilbert spaces*, Proc. Amer. Math. Soc. **142** (2014), no. 11, 3889–3895. MR 3251728
- [AL08] Akram Aldroubi and Casey Leonetti, *Non-uniform sampling and reconstruction from sampling sets with unknown jitter*, Sampl. Theory Signal Image Process. **7** (2008), no. 2, 187–195. MR 2451766 (2009k:94067)
- [Aro43] Nachman Aronszajn, *La théorie des noyaux reproduisants et ses applications. I*, Proc. Cambridge Philos. Soc. **39** (1943), 133–153. MR 0008639 (5,38e)
- [Aro48] ———, *Reproducing and pseudo-reproducing kernels and their application to the partial differential equations of physics*, Studies in partial differential equations. Technical report 5, preliminary note, Harvard University, Graduate School of Engineering., 1948. MR 0031663 (11,187b)
- [BJ02] Ola Bratteli and Palle Jorgensen, *Wavelets through a looking glass*, Applied and Numerical Harmonic Analysis, Birkhäuser Boston, Inc., Boston, MA, 2002. MR 1913212 (2003i:42001)
- [CFM⁺13] Peter G. Casazza, Matthew Fickus, Dustin G. Mixon, Jesse Peterson, and Ihar Smalyanau, *Every Hilbert space frame has a Naimark complement*, J. Math. Anal. Appl. **406** (2013), no. 1, 111–119. MR 3062405
- [CS02] Felipe Cucker and Steve Smale, *On the mathematical foundations of learning*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), no. 1, 1–49. MR 1864085 (2003a:68118)
- [DS88] Nelson Dunford and Jacob T. Schwartz, *Linear operators. Part II*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1988, Spectral theory. Selfadjoint operators in Hilbert space, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1963 original, A Wiley-Interscience Publication. MR 1009163 (90g:47001b)
- [HKL⁺14] S. Haeseler, M. Keller, D. Lenz, J. Masamune, and M. Schmidt, *Global properties of Dirichlet forms in terms of Green’s formula*, ArXiv e-prints (2014).

- [HN14] Haakan Hedenmalm and Pekka J. Nieminen, *The Gaussian free field and Hadamard's variational formula*, Probab. Theory Related Fields **159** (2014), no. 1-2, 61–73. MR 3201917
- [HQKL10] Minh Ha Quang, Sung Ha Kang, and Triet M. Le, *Image and video colorization using vector-valued reproducing kernel Hilbert spaces*, J. Math. Imaging Vision **37** (2010), no. 1, 49–65. MR 2607639 (2011k:94032)
- [JP10] Palle Jorgensen and Erin P.J. Pearse, *A Hilbert space approach to effective resistance metric*, Complex Anal. Oper. Theory **4** (2010), no. 4, 975–1013. MR 2735315 (2011j:05338)
- [JP13] ———, *A discrete Gauss-Green identity for unbounded Laplace operators, and the transience of random walks*, Israel J. Math. **196** (2013), no. 1, 113–160. MR 3096586
- [JPT14] Palle Jorgensen, Steen Pedersen, and Feng Tian, *Restrictions and extensions of semibounded operators*, Complex Anal. Oper. Theory **8** (2014), no. 3, 591–663. MR 3167762
- [JS09] Palle E. T. Jorgensen and Myung-Sin Song, *An extension of Wiener integration with the use of operator theory*, J. Math. Phys. **50** (2009), no. 10, 103502, 11. MR 2573134 (2011b:60214)
- [JS12] ———, *Comparison of discrete and continuous wavelet transforms*, Computational complexity. Vols. 1–6, Springer, New York, 2012, pp. 513–526. MR 3074509
- [JS13] ———, *Compactification of infinite graphs and sampling*, Sampl. Theory Signal Image Process. **12** (2013), no. 2-3, 139–158. MR 3285408
- [JT15a] P. Jorgensen and F. Tian, *Graph Laplacians and discrete reproducing kernel Hilbert spaces from restrictions*, ArXiv e-prints (2015).
- [JT15b] Palle Jorgensen and Feng Tian, *Discrete reproducing kernel Hilbert spaces: Sampling and distribution of Dirac-masses*, Journal of Machine Learning Research **16** (2015), 3079–3114.
- [JT15c] Palle Jorgensen and Feng Tian, *Infinite networks and variation of conductance functions in discrete Laplacians*, J. Math. Phys. **56** (2015), no. 4, 043506, 27. MR 3390972
- [KGD13] Samir Khan, Roger M. Goodall, and Roger Dixon, *Non-uniform sampling strategies for digital control*, Internat. J. Systems Sci. **44** (2013), no. 12, 2234–2254. MR 3172778
- [KH11] Sanjeev Kulkarni and Gilbert Harman, *An elementary introduction to statistical learning theory*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, NJ, 2011. MR 2908346
- [KZ96] Samuel Karlin and Zvi Ziegler, *Some inequalities of total positivity in pure and applied mathematics*, Total positivity and its applications (Jaca, 1994), Math. Appl., vol. 359, Kluwer Acad. Publ., Dordrecht, 1996, pp. 247–261. MR 1421605 (98d:15019)
- [Lan67] H. J. Landau, *Necessary density conditions for sampling and interpolation of certain entire functions*, Acta Math. **117** (1967), 37–52. MR 0222554 (36 #5604)
- [LB04] Yi Lin and Lawrence D. Brown, *Statistical properties of the method of regularization with periodic Gaussian reproducing kernel*, Ann. Statist. **32** (2004), no. 4, 1723–1743. MR 2089140 (2006a:62053)
- [LP11] Sneh Lata and Vern Paulsen, *The Feichtinger conjecture and reproducing kernel Hilbert spaces*, Indiana Univ. Math. J. **60** (2011), no. 4, 1303–1317. MR 2975345
- [MGG14] Adam Martinez, Anne Gelb, and Alexander Gutierrez, *Edge detection from non-uniform Fourier data using the convolutional gridding algorithm*, J. Sci. Comput. **61** (2014), no. 3, 490–512. MR 3268657
- [Nel57] Edward Nelson, *Kernel functions and eigenfunction expansions*, Duke Math. J. **25** (1957), 15–27. MR 0091442 (19,969f)
- [OST13] Kasso A. Okoudjou, Robert S. Strichartz, and Elizabeth K. Tuley, *Orthogonal polynomials on the Sierpinski gasket*, Constr. Approx. **37** (2013), no. 3, 311–340. MR 3054607
- [RW06] Carl Edward Rasmussen and Christopher K. I. Williams, *Gaussian processes for machine learning*, Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, 2006. MR 2514435 (2010i:68131)
- [SS01] Bernhard Scholkopf and Alexander J. Smola, *Learning with kernels: Support vector machines, regularization, optimization, and beyond (adaptive computation and machine learning)*, 1st ed., The MIT Press, 12 2001.

- [SS13] Oded Schramm and Scott Sheffield, *A contour line of the continuum Gaussian free field*, Probab. Theory Related Fields **157** (2013), no. 1-2, 47–80. MR 3101840
- [ST12] Robert S. Strichartz and Alexander Teplyaev, *Spectral analysis on infinite Sierpiński fractafolds*, J. Anal. Math. **116** (2012), 255–297. MR 2892621
- [STC04] John Shawe-Taylor and Nello Cristianini, *Kernel methods for pattern analysis*, Cambridge University Press, 2004.
- [Sto51] M. H. Stone, *On unbounded operators in Hilbert space*, J. Indian Math. Soc. (N.S.) **15** (1951), 155–192 (1952). MR 0052042 (14,565d)
- [Str10] Robert S. Strichartz, *Transformation of spectra of graph Laplacians*, Rocky Mountain J. Math. **40** (2010), no. 6, 2037–2062. MR 2764237 (2012c:05199)
- [SZ09] Steve Smale and Ding-Xuan Zhou, *Online learning with Markov sampling*, Anal. Appl. (Singap.) **7** (2009), no. 1, 87–113. MR 2488871 (2010i:60021)
- [Vul13] Mirjana Vuletić, *The Gaussian free field and strict plane partitions*, 25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2013), Discrete Math. Theor. Comput. Sci. Proc., AS, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013, pp. 1041–1052. MR 3091062
- [ZLP14] Lincheng Zhou, Xiangli Li, and Feng Pan, *Gradient-based iterative identification for Wiener nonlinear systems with non-uniform sampling*, Nonlinear Dynam. **76** (2014), no. 1, 627–634. MR 3189198
- [ZXZ12] Haizhang Zhang, Yuesheng Xu, and Qinghui Zhang, *Refinement of operator-valued reproducing kernels*, J. Mach. Learn. Res. **13** (2012), 91–136. MR 2913695

(PALLE E.T. JORGENSEN) DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF IOWA, IOWA CITY, IA 52242-1419, U.S.A.

E-mail address: `palle-jorgensen@uiowa.edu`

URL: `http://www.math.uiowa.edu/~jorgen/`

(FENG TIAN) DEPARTMENT OF MATHEMATICS, HAMPTON UNIVERSITY, HAMPTON, VA 23668, U.S.A.

E-mail address: `feng.tian@hamptonu.edu`